CATEGORIES OF HUMAN LEARNING

Contributors to this Volume

Norman H. Anderson
E. James Archer
G. E. Briggs
James Deese
W. K. Estes
Paul M. Fitts
Robert M. Gagné
David A. Grant
Howard H. Kendler
Tracy S. Kendler
Gregory A. Kimble
Arthur W. Melton
Leo Postman
Benton J. Underwood
Delos D. Wickens

CATEGORIES OF HUMAN LEARNING

Edited by

ARTHUR W. MELTON

DEPARTMENT OF PSYCHOLOGY
THE UNIVERSITY OF MICHIGAN
ANN ARBOR, MICHIGAN

1964

ACADEMIC PRESS New York · London

ACADEMIC PRESS INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS INC. (LONDON) LTD
Berkeley Square House, London W.1

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 63-14566

Fourth Printing, 1968

PRINTED IN THE UNITED STATES OF AMERICA

List of Contributors

Numbers in parentheses refer to the pages on which the contributors' articles begin

NORMAN H. ANDERSON (129), University of California, Los Angeles, California[1]

E. JAMES ARCHER (237), University of Wisconsin, Madison, Wisconsin[2]

G. E. BRIGGS (286), Ohio State University, Columbus, Ohio

JAMES DEESE (202), The Johns Hopkins University, Baltimore, Maryland

WILLIAM K. ESTES (89), Stanford University, Stanford, California

PAUL M. FITTS (243), The University of Michigan, Ann Arbor, Michigan[3]

ROBERT M. GAGNÉ (293), Princeton University, Princeton, New Jersey[4]

DAVID A. GRANT (1), University of Wisconsin, Madison, Wisconsin

HOWARD H. KENDLER (211), New York University, New York, New York[5]

TRACY S. KENDLER (318), Barnard College, New York, New York[6]

GREGORY A. KIMBLE (32), Duke University, Durham, North Carolina

ARTHUR W. MELTON (325), The University of Michigan, Ann Arbor, Michigan

LEO POSTMAN (145), University of California, Berkeley, California

BENTON J. UNDERWOOD (47), Northwestern University, Evanston, Illinois

DELOS D. WICKENS (79), Ohio State University, Columbus, Ohio

[1] Present Address University of California, San Diego, California
[2] Present Address University of Colorado Boulder, Colorado
[3] Deceased 2 May 1965
[4] Present Address University of California, Berkeley, California
[5] Present Address University of California, Santa Barbara, California
[6] Present Address University of California, Santa Barbara California

Preface

In this volume are recorded the papers of participants in a *Symposium on the Psychology of Human Learning* which was held at the University of Michigan, Ann Arbor, on January 31 and February 1, 1962 The symposium, and a two-day meeting of the participants following the symposium, were sponsored by the Office of Naval Research under Contract Nonr-1224(37) with the University of Michigan By agreement among the participants and for reasons that will become clear in a moment, the more descriptive title, *Categories of Human Learning*, was selected for this collection of printed papers

The division of the meeting of the participants into two parts—one public, and devoted to the presentation of prepared papers, and the other private, and devoted to issues raised in the papers as well as to some more general questions relating to critical issues in the psychology of human learning—was a deliberate plan to spread the immediate stimulating effects of the prepared papers to graduate students and staffs at nearby colleges and universities as well as to those located at the University of Michigan This purpose was well served, and my only regret was that the announcements and invitations to the symposium were not spread through out a geographical area with a radius much greater than the one chosen Even so, some 90 graduate students and faculty members from 15 different colleges, universities, and research laboratories in Michigan (other than Ann Arbor) and in Ohio attended the symposium They represented approximately 40 per cent of the audience at each symposium session The remainder of some 200 persons in the audience were from the University of Michigan

While the announced title of the symposium was *The Psychology of Human Learning*, the fine print of the announcement specified that "The theme of the Symposium is the interrelationship of different categories of human learning, with emphasis on (*a*) the formulation of statements about the categories that reflect the sophisticated methodology of contemporary laboratory studies, and (*b*) the relatedness of the different categories of human learning—whether in laboratory practice, in empirical generalizations about organismic or procedural variables, or in theory" This statement of the theme was, of course, a synoptic statement of the purpose of the symposium and conference, as stated in the instructions to the participants

The intent was to focus the attention of each of the seven principal

participants on the definitional and taxonomic issues that plague the psychology of human learning as well as the psychology of learning in general My reasons for believing that such an effort was needed, and would be worthwhile, are developed in some detail in my concluding comments on the symposium (see Chap VIII) Suffice it to say at this time that it seemed to me that these taxonomic issues should at least be faced as a prolegomena to any systematic organization of knowledge about human learning as well as to other ventures such as the topical breakdown of the science of human learning for the purpose of organizing a multi author handbook or the selection of particular 'kinds" of human learning as the topics of special symposia

It was not expected that the principal participants and their discussants would provide psychologists with a new taxonomy of human learning, with solutions to some of the perennial taxonomic problems of learning (e g, classical vs instrumental conditioning, concept learning vs problem solving), or with persuasive evidence that our taxonomic problems are phenotypic mirages readily dispelled by genotypic theory It was expected that the problems of coming to grips with these taxonomic issues would be clarified and that the conclusions, even the failures, of these experts in each of the selected 'kinds" of human learning would form an explicit point of departure for further thought and research on this fundamental systematic problem in the psychology of human learning These expectations have, in my opinion, been realized For this I am deeply indebted to each of the participants, even though the earnestness of their endeavors to face the issues posed for them and the insightfulness of their contributions came as no surprise

I am sure that some psychologists will wonder why the seven categories of human learning chosen for this symposium were not a different seven or were not eight, nine, or even more The seven were chosen because they seemed to represent the categories most commonly employed by investigators in thinking about and doing research on human learning and have become for this reason part of the traditional descriptive language of the science of human learning Admittedly, some of the categories might have been split into two categories, but this would have resulted in distribution of responsibility for considering the interrelations of the sub-categories Hind sight suggests that it would have been desirable to have a contribution on 'perceptual learning' but this was explicitly considered and rejected by me at the time the symposium was planned

As indicated above the symposium was followed by a two-day meeting of the participants during which they continued their discussions of the papers that had been presented, attempted to identify places in their arguments where the exposition could be clarified or the issues could be more

sharply drawn, etc. The papers that are printed in this volume reflect in many instances the sharpening effect of these discussions, because authors were encouraged to submit for publication a revision of the papers actually presented at the symposium if such revision better reflected their final positions with respect to the issues involved.

Another function of the post-symposium meeting was to consider the question of possible topics for subsequent symposia in the general area of human learning. In the first place, the concensus of the group was that symposia that stressed, or at least explicitly included, the definitional and taxonomic issues of the present symposium were needed for the improvement of communication among psychologists interested in human learning. This is not to say that another symposium of the same breadth of concern for the interrelations between "types" of learning as the present one should be held, but rather that the interrelatedness of one type of learning and every other—at least to some degree, or in some sub-classes of a category—should be recognized in structuring the special category-oriented symposia. For example, a symposium on "concept formation and utilization" should have someone who looks at the problem from the point of view of research on "voluntary instructed conditioning," from the point of view of rote learning, from the point of view of theories of skill learning, from the point of view of theories of short-term memory, etc. Similarly, a symposium on probability learning should have participants who look at that problem from the point of view of rote learning, short-term memory, concept learning, problem solving, etc. Of course, this is what some investigators do when they perform "transition experiments" (Underwood, p. 49), and this is what some theorists do. But the largely artificial boundaries between the "kinds" of learning, and the not so artificial limitations of the detailed knowledge of the literature about these various "types" of learning among specialists in each type, need to be forcibly bridged in the interests of promoting more comprehensive understanding of the interrelations of forms of human learning.

When the group turned to consideration of the most promising topics for the advancement of our understanding of human learning and the interrelations of types of learning, it was generally agreed that three topics deserved priority, although among these three there was no agreed-upon priority and none was thought necessary. First, there seemed to be a need for a full-scale consideration of the varieties of learning properly described as "conditioning." This should include animal conditioning as well as human conditioning because (*a*) the operations of classical, iostrumental, and operant cónditioning are those most readily applicable across different kinds of organisms, and (*b*) the consistency and continuity of definitional and taxonomic issues across classes of organisms is so essential to the

proper grounding of the science of human learning in the more general science of organismic behavior In addition to including definitional and interrelatedness analyses of what may be described as the basic categories of conditioned response learning, one of the important classes of contributions needed in this symposium was thought to be the systematic coverage of attempts to apply the conditioning paradigm or to test for some of the principal phenomena of conditioned response learning (e g , experimental extinction, spontaneous recovery, partial reinforcement effects, CS-UCS temporal relations) in other categories of human learning Important among these other categories would be rote verbal learning and memory, probability learning, concept learning, and the establishment of mental sets in a wide variety of human tasks involving the establishment of new performance capabilities or the utilization of previously established capabilities The selection of "conditioning" for such emphasis reflected a conviction (not necessarily unanimous nor of the same strength in all participants) that the concepts, operations, aod phenomena of conditioning represent at least one of the major anchors for a systematic description of the varieties of human learning

The second topic thought to be deserving of special emphasis was that of second-order habits (concepts, rules, sets) in all forms of human learning—their development (and loss) and utilization It seemed clear from the papers presented in the symposium that all 'kinds" of human learning were subject to the influence of such factors and that one of the possibly important sources of overlap and similarity of effects of independent variables in the varieties of human learning might well be their common involvement of these "second order" habits It was thought that here, for sure, would be a topic that required the participation of psychologists primarily concerned with perceptual learning, as well as with the other categories of human learning represented in the symposium

The third topic thought deserving of emphasis was that of mediating processes The postulation of mediating processes occurred in the discussion of every category of human learning and yet there seems to be no agreed-upon set of criteria for postulating mediating processes, no clear-cut set of defining operations for mediating processes, and no taxonomy of mediational concepts or mediating processes If, as seems to be the case, our theoretical descriptions of all varieties of human learning requires the postulation of mediating processes, it seemed to the group that one of the most significant contributions to the science of human learning should be a concerted effort to assess the empirical evidence for such processes, to devise and assess observable indices or correlates of such processes, and to put our conceptual language for such processes in order

These three topics were not the only ones discussed, but they reflect

adequately the flavor of the thinking of the group about the problems that must be faced in any attempt to put some greater orderliness into the conceptual framework of the psychology of human learning

In addition to my enduring indebtedness to all of the participants in this symposium for their contributions and for the supportative attitudes they expressed with respect to the purposes of the venture, I also have two further obligations to express The first is to Dr D D Smith, then Chief of the Personnel and Training Branch of the Office of Naval Research, who dogged my tracks for some months in an effort to persuade me to take on the responsibility for arranging a symposium on some problem in the area of human learning, and who, when I finally succumbed, gave complete support to the idea of having a symposium on the problem of the interrelatedness of different categories of human learning Without his persistent encouragement I would not have entered into the venture, and I would have been denied the satisfaction that comes from seeing a problem of enduring concern to me so ably and interestingly handled by one and all

I, and all members of the symposium, are also indebted to Miss Frances A Hill, who served voluntarily as secretary of the group during the two days of the conference following the symposium and provided me with an effective summary of the discussion and conclusions She also prepared the subject index for this book

ARTHUR W MELTON
The University of Michigan

December, 1963

Contents

The Centrality of Verbal Learning
Comments on Professor Underwood's Paper

DELOS D WICKENS

Probability Learning

WILLIAM K ESTES

An Evaluation of Stimulus Sampling Theory
Comments on Professor Estes' Paper

NORMAN H ANDERSON

Short Term Memory and Incidental Learning

LEO POSTMAN

Behavioral Effects of Instruction to Learn
Comments on Professor Postman's Paper

JAMES DEESE

The Concept of the Concept

HOWARD H KENDLER

On Verbalizations and Concepts
Comments on Professor Kendler's Paper

E JAMES ARCHER

Perceptual Motor Skill Learning

PAUL M FITTS

Classical and Operant Conditioning

DAVID A. GRANT
University of Wisconsin

All the topics of this symposium deal with learning in the sense that they deal with the establishment and strengthening of S—R connections or associations It is quite appropriate to start such a discussion or symposium with classical and operant conditioning for two reasons First there is some historical priority, and second, in classical and operant conditioning there are probably fewer S—R connections involved so that these experiments are, in a sense, simpler than most of the studies of human learning At the same time, we must note that no S—R connections exist in isolation, each is embedded in a complex matrix of behavior in any experiment, and when we isolate a particular S—R connection for logical consideration we can never really isolate it from its matrix And although we may consider conditioning to be simpler than most human learning in fact, there appear to be all sorts of processes associated with these apparently simple experiments which make their phenomena quite complicated Indeed, so highly efficient and developed are the verbal capacities of our human subjects (*S*s) that some of the fairly complicated forms of human learning in which verbal functioning is especially involved actually seem to give us simpler behavior, or at least behavior which is more regular, less variable, than that found in our typical conditioning experiments Also human learning occurs more readily and more rapidly in some experiments that tap verbal functioning than in experiments that do not For example a simple discrimination and discrimination reversal utilizing verbal responding occurs almost immediately and with almost perfect accuracy, whereas if the same experiment is done, requiring a differential conditioned eyelid response of the human *S*, both the initial differential conditioning and the subsequent reversal are much slower, much less perfect, and "contaminated" with other complexities which will not be apparent when the verbal response is used (Clark, Grant, & Levy, 1963)

The scheme of this paper is as follows *first*, I shall present a way of classifying classical conditioning experiments into four subclasses The four subclasses involve, successively less and less motivation from extrinsic sources and more and more dependence upon the conditioned stimulus

(CS) and unconditioned stimulus (UCS) alone *Second*, a tridimensional classification of operant or instrumental experiments is outlined that depends upon cues, nature of the reinforcement, and nature of the response demanded of the *S* The two sets of subclassifications are necessarily operational and incompletely operational at that, rather than functional, so that the *third* section of the paper points out some of the further operations and functional relations that set limits to the operational classification Up to this point I shall not have distinguished sharply between human and animal conditioning Then in the *fourth* section I shall discuss briefly human classical and operant conditioning, both the simple forms and more complex varieties that parallel in some respects the basic varieties of human learning that form the subject matter of this symposium Considerations noted here lead to the *fifth* section which presents some general conclusions about our work on human learning, some of its complexities, and some of the interrelationships between different forms of human learning So we turn, then, to the classification of conditioning experiments

Although early workers did not make the distinction, it has proved useful for over twenty years to separate conditioning experiments into two broad classes classical and instrumental conditioning The Pavlovian experiment in which dogs learned to salivate to a previously neutral stimulus which had been paired repeatedly with food presentations has been taken to be the prototype of classical conditioning Likewise the Thorndikian experiment in which hungry cats learned to manipulate latches, etc in order to escape from a puzzle box and obtain food has been the prototype for instrumental conditioning or learning The Skinnerian lever pressing by rats in order to obtain food is also frequently given as the prototype for instrumental conditioning, but the free operant presents special problems, so that it may be best to consider it as a class by itself

Attempts to distinguish between classical and instrumental learning in terms of what is learned and how it is learned break down under careful analysis, and psychologists have resorted to operational definitions to separate the two classes of conditioning experiments Thus Kimble (1961, p 44) states

> The basic distinction between classical and instrumental conditioning procedures is in terms of the consequences of the conditioned response In classical conditioning, the sequence of events is independent of the subject s behavior In instrumental conditioning by contrast rewards and punishments are made to occur as a consequence of the learner s response or failure to respond

This definition serves quite well, but some precision may be gained by stating that in instrumental conditioning the *availability* of reward or

punishment is contingent upon the S's behavior whereas in classical conditioning it is not

Within the broad framework provided by these definitions the originality of numerous experimenters has led to tremendous variations in procedures In some instances these variations in procedure tax the ingenuity of anyone attempting to classify the kind of learning involved Nevertheless valuable classificatory schemes have been provided by Hilgard and Marquis (1940), Konorski (1948), Razran (1961a, 1961b), and Kimble (1961) Kimble's revision of the Hilgard and Marquis classic serves as an admirable point of departure from which to elaborate subclassifications of classical and instrumental conditioning

Subclasses of Classical and Instrumental Conditioning

Subclasses of Classical Conditioning

In some respects subclassification of classical conditioning poses more problems than subclassification of instrumental conditioning Nevertheless it will be useful to distinguish between four different kinds of classical conditioning These will be designated Pavlovian A, Pavlovian B, Anticipatory Instructed, and Sensory Preconditioning Although S's responses and indeed his learning may be very similar in these four subclasses of classical conditioning the experimenter s operations are quite distinct

Pavlovian A conditioning—This kind of conditioning is exemplified by the prototype of classical conditioning In the original Pavlovian experiment the sound of a bell was repeatedly presented shortly before food was given the dog Originally the sound of the bell elicited orientation reactions such as looking toward the sound source and pricking up the ears When the food was presented it was ingested, eliciting the usual alimentary re flexes including buccal salivation After repeated paired presentations of bell and food the sound of the bell came to evoke salivation and orienta tion movements toward the feeding cups (Pavlov, 1927, p 22) These arrangements are diagrammed in Fig 1 The sound of the bell is called the

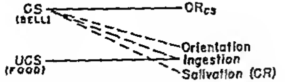

FIG 1 Diagram for Pavlovian A classical conditioning Pavlov selected salivation a single component of the complex of responses forming the UCR for study Por tions of the complex e g ingestion are absent in the CR The CS has a signaling function Performance of the UCR and subsequent conditioning are dependent upon food deprivation

conditioned stimulus (CS), the response to the bell is called the orienting reflex (OR), the food is called the unconditioned stimulus (UCS) and ingestion and salivation, etc , are called the unconditioned response (UCR) Note that buccal stimulation by the UCS requires ingestion, an instrumental act by *S* The new functional or learned connection is called the conditioned response (CR)

In this experiment many events occurred which Pavlov did not choose to measure In the complex of responses to the UCS Pavlov chose to record only salivation in most of his experiments Approach and ingestion were certainly responses to the UCS (food) and also could be noted in the CR But it is noteworthy that the CR did not include approach to and attempts to feed upon the bell or other source of the CS (Zener, 1937) The approach and orientation movements were directed to the food source These facts indicate that the CS does not *substitute* for the UCS as is so frequently stated Pavlov states that the CS serves as a signal that the food is about to be presented This position is also taken by Schlosberg (Kimble, 1961, p 99) and seems to be a more accurate way of designating the function of the CS

An important feature of Pavlovian A conditioning is the consummatory response, in this case, ingestion of the food Although Pavlov s dogs had doubtless formed pre-experimental conditioned salivary responses to the sight and odor of food they did, in fact, ingest the food in his experiments, and they were brought to the experimental situation after some hours of food deprivation It is known that conditioned salivary responses are hard to form in animals that are not hungry and, if formed, are hard to elicit from a satiated dog It follows then that results of Pavlovian A conditioning must always be related to the motivational state of the organism during acquisition and testing of the CR

Pavlovian B conditioning —This subclass of classical conditioning could well be called Watsonian conditioning after the Watson and Rayner (1920) experiment conditioning fear responses in Albert, but Pavlov has priority The reference experiment for Pavlovian B conditioning might be that in which an animal is given repeated injections of morphine The UCR to morphine involves severe nausea, profuse secretion of saliva, vomiting, and then profound sleep After repeated daily injections Pavlov's dogs would show severe nausea and profuse secretion of saliva at the first touch of the experimenter (Pavlov, 1927, pp 35–36) In Pavlovian B conditioning stimulation by the UCS is not contingent on *S*'s instrumental acts, and hence there is less dependence upon the motivational state of the organism, and the CS appears to act as a partial substitute for the UCS Furthermore, the UCS elicits the complete UCR in Pavlovian B conditioning whereas in Pavlovian A conditioning the organism emits the UCR of approaching

and ingesting the food Although Pavlovian B conditioning may be less affected by the motivational state of the organism the experimenter cannot ignore the presence of concurrent antagonistic responses For example Watson and Rayner (1920) reported that they could not evoke conditioned fear responses as long as Albert had his thumb in his mouth That the organism learns instrumental acts outside the paradigm for Pavlovian B conditioning is indicated by the fact that Albert evidently learned to put his thumb in his mouth when a frightening stimulus was presented Phenomena of this sort have not been investigated to any great extent but it is obvious that ignoring them can lead to some confusion in comparing results in different experiments

The diagram for Pavlovian B classical conditioning is shown in Fig 2

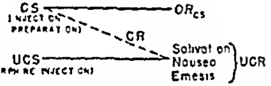

FIG 2 Diagram for Pavlovian B classical conditioning The CS substitutes to a great extent for the UCS producing most if not all of the UCR Conditioning and performance are relatively independent of motivation but probably dependent upon concurrent antagonistic responses e g feeding

The CS constitutes the preparations for the injection and insertion of the needle There will originally be an OR to the CS The UCS is the morphine injection itself a very complex physiological stimulus The UCR consists of salivation nausea vomiting and other physiological changes Many of these are readily seen as components of the CR which will be evoked by the preparations for the injection after repeated daily morphine injections A great deal of interoceptive conditioning (Bykov, 1957 Razran 1961b) and autonomic conditioning (Kimble 1961) apparently follows the Pavlovian B paradigm

Anticipatory instructed conditioning—Following Ivanov Smolensky (1933) many Russian investigators have formed what they, quite legitimately, I think, call CRs based upon the voluntary squeezing of a rubber bulb in response to a sound stimulus The sound stimulus is preceded regularly by a light or some other neutral stimulus and the *Ss* will learn to respond to the neutral stimulus In effect it is a reaction time experiment with a constant fore period signaled by the onset of the CS Some will doubtless balk at calling this conditioning It is however similar to transferring stimulus control of an operant from one S^D to another S^D An excellent reference experiment for anticipatory instructed conditioning is the

establishment of conditioned eyelid responses based on voluntary reinforcement by Marquis and Porter (1939) The diagram for Anticipatory Instructed Conditioning is shown in Fig 3 The CS was a change in illumina-

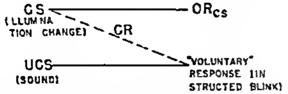

FIG 3 Diagram for Anticipatory Instructed Conditioning The CR appears as an anticipatory blink before the sound (UCS) Marquis and Porter (1939) found best conditioning when the instructions were to blink quickly and when the CS was sufficiently peripheral to attract little attention

tion The *Ss* were instructed to blink as rapidly as possible when they heard a sound With a weak peripheral CS 6 out of 10 *Ss* gave 59% to 92% anticipatory eyelid responses in 50 paired presentations of the light and sound The *Ss* were unable to verbalize the relationship between the change in illumination and the other stimuli of the experiment Generally speaking, extinction was quite rapid following Marquis and Porter's procedures It should be noted that of the 4 *Ss* who failed to show appreciable conditioning 2 spontaneously described their attempts to inhibit blinking to the light, whereas the other 2, although they did not verbalize an inhibitory attitude, showed long reaction times The form of the resulting CR was quite different from that obtained using the reflex wink to an air puff as the UCR This eyelid CR closely resembled the voluntary response described by Spence and his co-workers (Spence & Taylor, 1951, Spence & Ross, 1959), i e, the response was a complete closure with rapid recruitment and prolonged duration

Although Russian investigators have utilized this CR a great deal and have investigated its characteristics extensively (The Israel Program for Scientific Translations, 1960) these investigations are not well known in this country, and very little work has been done with this kind of experimental procedure It is evident, however, from the report of Marquis and Porter (1939) that this type of conditioning is heavily dependent upon stimulus arrangements instructions, and attitude of the *S*

In anticipatory instructed conditioning the CS appears to trigger a preset reaction The response might be viewed as the false reaction to the ready signal in a reaction time experiment It seems apparent that the CS may signal the occurrence of the UCS or may substitute for it, depending upon idiosyncrasies of the *S* The sequence of events in the conditioning

situation will therefore depend upon the *S*'s interpretation of the *E*'s instructions and the *S*'s own self instruction With peripheral, scarcely noticed, conditioned stimuli and delicate time relations in the response, the *S* may be unaware of the fact that he is responding to the "wrong" stimulus In the event that the *S* does not know that he has responded to the CS he may presumably interpret this as a "correct" or an "incorrect" response on his part Properly instructed and properly motivated he will presumably try to increase "correct" responses and decrease "*incorrect* responses In anticipatory instructed conditioning the definition of "correct" and "incorrect" are usually left ambiguous, the *E* is at the mercy of the *S*'s self-instructions

Interpretation of instructions is important in other conditioning and learning experiments where the *S* believes that his responses are "voluntary" This can account for some of the anomalies found in hand-withdrawal conditioning It is amusing but also instructive to go back to some of the earlier reports of hand withdrawal conditioning when extensive verbal reports were taken from the *S*s (e g , Hamel, 1919, Schilder, 1929) Here we find some *S*s who had to fight themselves to keep their hands down on the shocking grid in order not to spoil the *E*'s results, whereas other *S*s, as soon as they understood the relationship between the CS and the shock, withdrew their hands immediately, for to do otherwise would have been ridiculous to them In sum, the care with which instructions are given human *S*s in conditioning experiments can scarcely be overemphasized Leaving matters up to the *S* results in experiments where no two *S*s give the same sort of responses The results are more likely to be confusing than enlightening The enforced brevity of current reporting of psychological experiments puts a heavy but legitimate burden on the *E* It is his responsibility to get his experiment under control

Sensory preconditioning—A fourth subclass of classical conditioning is well represented by the sensory preconditioning experiment of Brogden (1939) *This is diagrammed in Fig 4 In the preconditioning phase there*

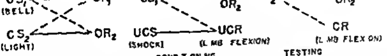

Fig 4 Diagram for Sensory Preconditioning The preconditioning phase produces sufficient *substitution potential between* CS_1 and CS_2 so that in testing phase CS_2 will evoke the CR previously conditioned to CS_1 only in conditioning phase

were 200 paired presentations of two conditioned stimuli During the conditioning phase, 20 daily, paired presentations of one CS with the shock to the left forelimb were presented until the criterion of 100% CRs was reached Then, in the testing phase, the other CS was presented 20 trials per day until the forelimb flexion response disappeared Eight out of 10 dogs produced forelimb flexion to the second CS in the testing phase and required from one to seven days to extinguish the response to the second stimulus The preconditioning produced effective stimulus substitution and is of considerable theoretical interest because of the motivational features or lack of them in the experiment It is by no means certain whether the preconditioning is S—S conditioning or S—R conditioning with chaining through the ORs to the two conditioned stimuli The CS-UCS interval in the conditioning phase was 2 sec which is quite adequate for chaining as has been demonstrated by Wickens and his co workers (e g , Wickens, 1959) There is evidently a good deal of published Russian research on sensory preconditioning which they call associative conditioning (cf Razran, 1961b)

Some of the American work on sensory preconditioning has not been as successful as the original Brogden experiment (cf Seidel, 1959) It is interesting in view of Wicken's results on compound conditioning (e g , Wickens, 1959) to conjecture that if the CS_2 preceded the CS_1 in the preconditioning phase then more preconditioning would be demonstrated in the testing phase of the sensory preconditioning experiment Although a good deal of sporadic work has been done on compound and chained conditioning, both in this country and in Russia (cf references listed by Wickens, 1959, Razran, 1961b), the events and interpretations in the more complicated experiments have not been worked out in any great detail

Subclasses of Instrumental Conditioning Experiments

In an instrumental conditioning experiment, contingent upon the *S*'s responses, other stimuli are made available to him, and the *S*'s interaction with these stimuli is the basis for the change in rate of emission of the original response Skinner (1937) has diagrammed the instrumental learning situation as shown in Fig 5 The subject emits a response, R_0, to un-

$$s \text{———} R_0 \longrightarrow S_1 \text{———} (R_1)$$

Fig 5 Instrumental conditioning according to Skinner (1937) The *S*, presumably on the basis of an internal stimulus (S) emits a response R_0, contingent upon which stimuli S_1, are made available The S_1——(R_1) complex constitutes the reinforcement. Learning is dependent upon the $R_0 \longrightarrow S_1$ contingency and is reflected in the rate of emission of R_0, the operant.

known stimuli, s Contingent upon the performance of R_0 other stimuli, S_1 are made available to the subject The S_1 and responses to S_1, (R_1), constitute the reinforcement Learning is dependent upon the $R_0 \rightarrow S_1$ contingency and leads to an increase in the rate or probability of emission of the original response, R_0

The extremely simple paradigm of Fig 5 admits to many variations in experimental arrangements Attention will be restricted to those in which only a single type of reinforcer is used Hilgard and Marquis (1940) distinguish four types of instrumental conditioning experiments reward training escape training avoidance training, and secondary reward training Essentially the same are mentioned by Woodworth and Schlosberg (1954, Ch 19) To these Kimble (1961) adds omission training and punishment training In further discussion it will be necessary to exclude secondary reward training as beyond the scope of this treatment As a first step, the reference experiments for each form of instrumental conditioning will be described

Reward training—The reference experiment for reward training might well be the free operant conditioning experiment of Skinner (1938) in which a rat is placed in a soundproof box containing a small lever which when pressed, automatically deposited a pellet of food in a delivery cup After a hungry rat has been placed in this situation his rate of lever pressing will usually increase until it reaches a steady value

Escape training—A good reference experiment for escape training was carried out by Sheffield and Temmer (1950) Rats were placed on a charged grill in an enclosed runway The floor of half of the runway consisted of the charged grill and the floor of the other half was smooth, providing a 'safe' escape compartment, with a sliding door in between to prevent retracing Speed of locomotion toward the safe compartment was used as the response measure

Reward and escape training performance serve operationally to define reward and punishment To parallel Thorndike's (1933) definition of satisfying and annoying states of affairs, the *S* will not avoid a reward and will often act so as to maintain or renew it A punishment involves a state of affairs which the *S* does nothing to preserve and usually acts so as to terminate

Avoidance conditioning—The conditioning phase of the Brogden sensory preconditioning experiment can serve as the reference experiment for avoidance conditioning The left forelimb of the dog rests on an electrode which is electrically charged upon the termination of a 2 sec CS On the first few trials the dog will receive the shock each time Eventually a conditioned anticipatory response involving struggling and lifting the forelimb from the

forcement When these three independent ways of dividing the subclasses of instrumental conditioning are simultaneously applied the results are as shown in the first six lines of Table 1 Each of the six varieties of instrumental conditioning fits neatly and unambiguously into a line of the table, dependent upon whether there is a cue to the reinforcement, the positive or negative nature of the response, and the nature of the reinforcement

TABLE 1

A Classification of Types of Instrumental Learning Experiments

Instrumental types	Cue to impending reinforcement	Response	Reinforcement
Reward Training	No	Positive	Reward
Escape Training	No	Positive	Punishment
Avoidance Conditioning	Yes	Positive	Punishment
Discriminated Operant	Yes	Positive	Reward
Omission Training	No	Negative	Reward
Punishment (or Passive Avoidance Conditioning)	No	Negative	Punishment
Discriminated Omission Training	Yes	Negative	Reward
Discriminated Punishment Training	Yes	Negative	Punishment

Arrangement of the six types of instrumental conditioning into such a tridimensional scheme pays an interesting dividend It will be noted that independent combinations of two possibilities as to the cue for reinforcement, two possibilities as to the response, and two possibilities as to the nature of the reinforcement yield, totally, eight possible experimental reference types Only six were described by Kimble It is instructive to fill out the two remaining possibilities in the table and to see if they represent plausible experiments

If a cue to the imminence of reinforcement is given and the learned response is to omit a normally performed act in order to obtain a reward we have the seventh line of the table which has been designated *Discriminated Omission Training* If a cue is given to the imminence of reinforce ment, and the appropriate response is omission or inhibition or a normally performed act in order to avoid punishment we have the eighth line of the table which has been designated *Discriminated Punishment Training* To my knowledge, psychologists have not performed these experiments formally, but the situation envisaged by the paradigm seems to occur as a commonplace in real life Discriminated omission training is represented

by the remark of the cynic that men seek the appearance of virtue rather than virtue itself. Suppose, for example, that a parent agrees with his daughter to reward her with money, a watch, or some other valuable consideration if she stops smoking, drinking, swearing, or some such behavior. The young lady, who may be maintaining her behavior at a high operant level, may interpret this bargain to mean that the undesirable behavior must not occur at times or in places where the parent may learn of it. The result may be a well-discriminated omission of the behavior in the presence of the parent or likely informants of the parent.

Discriminated punishment training is easily exemplified by one's colleagues who may have colorful or even lurid speech in informal social contacts but whose undergraduate lectures are models of propriety. The cynical expression is, "Avoid the appearance of evil." Training animals to desist from undesirable behavior is a rich field of examples for discriminated punishment training. I have had the personally unrewarding experience of trying to train numerous generations of cats not to scratch our furniture. The carefully conducted training procedure usually consists of my throwing a paperback book or whatever is handy at the cat when he approaches the furniture with evil intent or is engaged in scratching. Not only does our scratched up furniture give mute evidence for the fact that the animals have developed a discriminated punishment pattern, but also, if I slip into the room upon hearing a cat scratch the furniture I invariably find that the cat is indeed scratching and at the same time looking around the room to see if I am coming.

In spite of the practical importance of these learning patterns in discriminated omission training and discriminated punishment training these experiments seem to have been left severely alone by psychologists. To some extent discriminated omission training has been carried out by Skinner and his associates in DRL schedules where the *S* is rewarded for a low operant rate. Nevertheless, in view of the tremendous diversification of instrumental learning experiments it is somewhat surprising not to find any readily at hand which conform to these patterns. They seem well worth investigating.

This particular classificatory scheme is not proposed as the ideal scheme or even a complete scheme. It seems, however, to be serviceable and relatively unambiguous as a framework within which instrumental conditioning experiments may be placed. It would, perhaps, be more desirable to classify the experiments in terms of more fundamental psychological operations involved. It will be shown, moreover, that in view of the variety and complexity of experimental arrangements and the limitations in measurement of responses that the classification given above represents only the beginning of a lengthy and difficult task.

electrode will occur before the shock is given Finally, a regular, precise, and adaptive flexion of the forelimb will occur shortly before the electrode is charged At this stage of learning the dog will rarely receive a shock, and conditioned flexion will be maintained with a very small number of reinforcements

Discriminated operant conditioning—It is useful to differentiate discriminated operant conditioning from reward training or simple operant conditioning The reference experiment is a slight elaboration of the experiment for reward training An additional stimulus source is added to the rat's box, say an electric lamp Pellets are delivered by the apparatus when the lever is depressed if, and only if, the lamp is lit The rat soon learns to depress the lever only when the lamp is lit, and not to depress the lever when the lamp is not lit The operant is said to come under stimulus control, and the differential stimulus, the light, is designated by Skinner as S^D By clever arrangements of the contingencies between reinforcers and discriminatory stimuli, all sorts of remarkable behavior may be obtained (Ferster & Skinner, 1957)

Omission training—A reference experiment for omission training is that of Konorski (1948, p 226ff) A classical salivary CR to the sound of the metronome was occasionally evoked in combination with passive flexion of the dog's leg by the experimenter, and this combination of stimulation was never reinforced with food Very soon, it was found that concurrent raising of the leg with the sound of the metronome resulted in inhibition of the salivary flow In addition to this, at about the same time the animal began to resist the passive flexion of the leg by actively extending it Finally the extension movement became quite strong so that it could easily be recorded by means of a treadle arrangement The dog had learned not only to avoid the flexion movement associated with non reward but to make the antagonistic response For purposes of this discussion the antagonistic response need not be overt The situation roughly parallels that of rewarding a child when he restrains himself from some specified aggressive or destructive activity, but the use of rewards to make non verbal animals omit responses is likely to prove difficult without Konorski's "tutoring" procedure

Punishment training or passive avoidance—Another form of instrumental conditioning that Kimble derived from Konorski's valuable book was punishment training The neat procedure used by McCleary (1961) will serve as the reference experiment Hungry cats, placed in a two-compartment box, were trained to feed in the second compartment whenever a guillotine door was briefly raised After initial training had brought the feeding response to a high probability level the cat was shocked as it started

to feed After two such shocks McCleary found that normal cats would no longer feed or even approach the door of the feeding compartment when the door opened This passive avoidance was maintained in spite of several days of food deprivation

A Classificatory Scheme

In spite of the ingenuity of experimenting psychologists most instrumental conditioning experiments can be fitted into the six categories previously described Even so, there are numerous variations in procedure which will continue to defy any simple classificatory schema that may be proposed The problem still remains of trying to work out the relationships among the six categories that have been outlined Perhaps the most frustrating efforts to interrelate these subclasses of instrumental learning will be attempts based on what is learned Here the intentions and observations of the *E*, the pretraining of the *S*, the specific kinds of rewards and punishments, the rate of satiation, etc, all interact to produce a bewildering sequence of behavior, some classically conditioned, some instrumentally conditioned, most of which is ignored by the psychologist who has done the experiment

It turns out, however, that the different forms of instrumental conditioning do have a very simple tridimensional relationship if they are viewed in terms of *E*'s operations and the response reinforcement contingencies First of all, it will be noted that in some experiments rewards are used and in other experiments punishments are used Typically the learned response involves attaining the reward or avoiding the punishment Thus reward training, discriminated operant conditioning and omission training all involve reward as the reinforcement Escape training avoidance conditioning, and passive avoidance all involve punishment Furthermore, the experiments can be dichotomized in terms of whether the response of the *S* is the performance of a positive response or the omission or inhibition of a response which already had an operant strength greater than zero Thus, reward training, escape training, avoidance conditioning and discriminated operant conditioning all involve the performance of a positive response in order to attain the reward or to escape the punishment Omission training and passive avoidance conditioning both involve omission or inhibition of a response in order to achieve the reward or to avoid the punishment Finally, the experiments can be dichotomized in terms of whether or not a cue is given as to the time that the reinforcement will be available Thus in avoidance conditioning and discriminated operant conditioning a cue is available when the punishment or reward, respectively, is about to be presented In reward training, escape training, omission training and passive avoidance there is no special S^D or cue to the imminence of the rein

Special Problems of Equivalence of Stimuli, Responses, and Reinforcement in a Classificatory Scheme

Ideally a classificatory scheme for conditioning and learning ought to separate the different kinds of experiments into homogeneous classes so that the experiments within each class result in similar behavioral laws and phenomena, and experiments belonging to different classes result in different laws and phenomena The foregoing classification is in terms of some of the experimenter's operations, but within each class there is additional important variation in these operations, so that experiments classified alike do not, in fact, always yield similar behavioral laws A brief sampling of some of the important differences in identically classified experiments will be presented in terms of the stimuli, the responses, and the reinforcement problems

Stimuli

Experiments that are identical in terms of the limited classificatory scheme proposed above often have entirely different kinds of stimuli and sets of manipulanda and their associated stimuli owing to differences in apparatus Furthermore, because of the *S*s' previous experience in similar situations, learning sets (Harlow, 1949), observing responses (Wyckoff, 1952), special temporal relations among the stimuli that permit mediation and chaining (e g , Shipley, 1935, Wickens, 1959, Dicken, 1961), may differ tremendously from experiment to experiment In addition to these stimulus problems there are also the variations in drive stimuli and stimuli from concurrent activities All of these serve to produce functional differences between experiments which otherwise fit the same classificatory category

Response Definition

Classical conditioning experiments —As stated earlier, real refinements of the experimental arrangement so as to get meaningful conditioning results can be achieved by defining the conditions of the experiment so as to rule out irrelevant forms of behavior, and apparent refinements can result from a deliberate choice among responses available for observation As a simple example, Pavlov carried out his reference experiments under well-controlled physical conditions, excluding ambient noises, vibration, using healthy, alert, and hungry dogs, etc The animals were well isolated from the *E* and his controlling equipment, and Pavlov deliberately chose to record and measure salivation and, for the most part, to ignore other aspects of the dog s behavior

An excellent example of the consequences of inadequate control of extremely relevant concurrent activity and learning is seen in Shearn's (1961) review of the literature on cardiac conditioning Any given response system is intimately related to other response systems This is especially the case with vegetative response systems, such as those involved in heart rate When an electric shock is used as the UCS the gross conditioned change in the heart rate is an acceleration With recent advances in instrumentation it has been possible to record momentary heart rate continuously And when momentary heart rate is recorded cardiac conditioning becomes quite perplexing Sometimes the CR is an acceleration and sometimes a deceleration of the rate Some, at least, of the contradictory results of cardiac conditioning experiments, which led some investigators to doubt if cardiac conditioning really took place, apparently were caused by ignoring the conditioning of concurrent responses which influence the heart rate A typical initial response to electric shock is a sharp inspiration followed by a slower expiration This is readily conditioned Since the momentary heart rate increases with inspiration and decreases with expiration, one can find apparent conditioned acceleration or conditioned deceleration of the heart rate depending upon the temporal arrangements in the experiment Westcott and Huttenlocher (1961) brought some degree of order into cardiac conditioning by the simple expedient of having the *S*s control their breathing Of course the whole story has not been worked out as yet Owing to the complexity of the cardiac response system disentangling it from the respiratory reflex represents only the beginning of wisdom

A more detailed illustration of the problems involved in arranging the conditioning experiment and defining the CR occurs in the history of eyelid conditioning It is a history of progressive refinement in method extending almost 40 years Cason s pioneer work on the conditioned eyelid reaction (1922) was, as is usually the case, carried out with rather crude instrumentation His equipment would record any large blink, and test trials on which the UCS was omitted were necessary in order to evaluate conditioning Graphic records were not provided The CS—UCS interval was not optimal, and 500 to 1000 trials were usually required in order to obtain conditioning Hilgard s (1931) elegant eyelid conditioning experiments were carried out with the Dodge photochronograph which provided a photographic record of each trial The eyelid CR was revealed to be an anticipatory blink occurring before the UCR so that it was easily differentiated from the UCR and also from the primary reflex (alpha response) to the light Hence no test trials were needed in order to follow the course of conditioning The temporal separation of the reflex to light and the CR is shown in Hilgard s distribution of eyelid response latencies in Fig 6

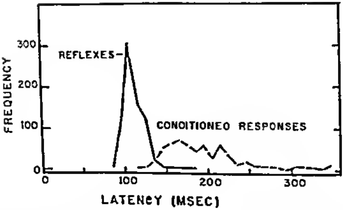

FIG 6 Distribution of the latencies of eyelid responses from Hilgard s (1931) experiment showing the separation of the primary reflex to light from the CRs on the basis of latency

Low magnitude CRs could be measured on the photographic records and typically occurred with considerable frequency early in conditioning This led Cason (1934) to complain that Hilgard was paying attention to minor twitches in the eyelid, but the superiority of the photographic method was quite apparent in the rapid and orderly conditioning of the eyelid

The photographic recording technique, however, also had its problems It was convenient to carry out the experiment in relative darkness so that, except in the Weber and Wendt experiment (1942), the *S* became progres sively dark adapted In some instances, once conditioned, the eyelid CR was hard to extinguish or even increased in frequency during the extinction trials Another striking feature was the fact that the early anticipatory blinks typically involved a reopening of the eye just in time to receive the air puff which was used as the UCS Although learned responses are not always adaptive, a maladaptive or nonadaptive response is likely to attract one s attention I thought that the peculiarity was an artifact of pseudo- conditioning (Kimble, 1961, pp 60–64) and pursued this line of investiga tion At Hilgard s suggestion, I added a crucial control group which even tually led to the discovery that there was a secondary eyelid response to light that became progressively sensitized during dark adaptation I called this the beta response, and Grant and Norris (1947) were able to show that with dark adapted *S*s a tremendous number of beta responses occur within the latency range of conditioned eyelid responses The point is illus- trated by comparing the latency distribution of eyelid responses in Figs 6 and 7 which represent the Hilgard and the Grant and Norris results, respectively By eliminating dark adaptation or by using low intensity con ditioned stimuli it was possible effectively to eliminate the beta response This had two effects it eliminated from the data a response which was

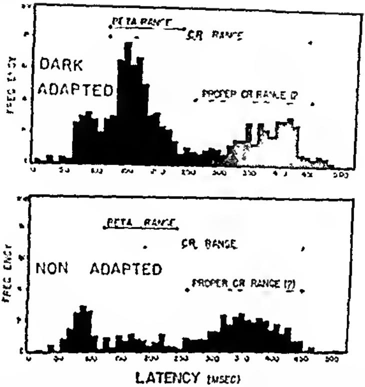

FIG 7 Latency distributions of eyelid responses evoked by light in an eyelid conditioning experiment in which half of the *Ss* were dark adapted and half were not (Grant & Norris 1947) The beta response, a secondary reflex to light, becomes sensitized during progressive dark adaptation and its increase in frequency resembles conditioning so that it was often confused with the CR in earlier experiments where the Dodge Photochronograph was used These responses can be largely eliminated by using a weak stimulus or by avoiding dark adaptation of the *S*

easily confused with the CR, and it prevented the refractory phase of this response, which normally precedes the CR, from interfering with the process of conditioning When the secondary eyelid response was controlled, orderly acquisition and extinction results were regularly obtained as is shown in Fig 8 which presents unpublished data from an experiment by Grant and Schneider (1948)

Although eliminating the beta response greatly improved matters, Spence and his co-workers (Spence & Taylor, 1951, Spence & Ross, 1959) were able further to refine the definition of the conditioned eyelid response Some *Ss* in eyelid conditioning experiments show a very abrupt acquisition curve and indeed may show this after one reinforcement These *Ss* also give slow extinction curves In examining their records, Spence discovered that these anomalous *Ss* were giving what appeared to be the voluntary responses of the Marquis and Porter (1939) experiment, i e , short latency, rapid recruit-

produce high response rates, such as fixed ratio schedules, are used, the performance of these two pigeons will be incomparably different The first pigeon that is addicted to wandering around the cage may be said to have poor response topography as compared with the pigeon who stays in the neighborhood of the button and vending equipment Usually some pretraining is carried out in order to shape the *S*'s response into a satisfactory topography With rats, many *E*s customarily smear a little food on the lever in order to center the animal's behavior on the manipulanda that are important for the experiment The amount of reward vended upon each occasion of reinforcement is also important here for very obvious reasons In any event, comparability between experiments is reduced if the operants in question are markedly dissimilar Since the effects of the pretraining itself may be unknown it would seem to be desirable to design the equipment so that the manipulanda lead rather directly to an operant of fairly uniform good topography The design of manipulanda to bring this about and indeed the techniques of pretraining themselves are arts which though widely known are not too well understood and certainly are not communicated in the literature in an orderly fashion

Even greater problems of achieving comparability arise when different types of equipment are used to carry out equivalent experiments For example, avoidance conditioning might be carried out in a shuttle-box where the instrumental response is that of leaving one compartment and going to another in order to avoid shock when a cue stimulus is presented Or the equivalent experiment could be carried out in a Skinner Box where pressing a lever is the instrumental act required in order to avoid a shock which would otherwise be given shortly after presentation of a cue stimulus Ordinarily the shuttle-box will produce faster learning than the Skinner Box in this situation Why functionally equivalent instrumental acts differ so much in learning difficulty is not entirely clear, although one could readily speculate that locomotion is a more "natural" response to avoid a punishing state of affairs Such an interpretation, although persuasive, begs many questions and has limited scientific value But some sort of equivalence operations must be established, or we will remain with one set of laws for the shuttle box and another set of laws for the Skinner Box Comparable problems arise if the behaviors of different species are to be compared For example rats and pigeons may be allowed to move about freely in the Skinner Box, but rhesus monkeys do better if they are partially restrained Although the need for restraint is an empirical fact and is understood on this basis, the role of restraint is not clear and will doubtless be found to produce its own artifacts

A good example of how attention to the detailed topography of a response may clarify the interpretation of the experiment is given by Kimble (1961,

Ch 4) Brogden, Lipman, and Culler (1938) had obtained faster conditioning of a running response of guinea pigs with avoidance training than with a conditioning procedure involving an unavoidable shock Some authors had interpreted this finding as a refutation of the contiguity principle in learning Sheffield was not satisfied with this interpretation and essentially replicated the experiment, obtaining the same finding (1948) Sheffield, however, noting that the unavoidable shock was frequently administered while the guinea pigs were running noted the precise effect of the shock on the *S*'s behavior On some trials the guinea pigs would continue to run and on other trials would jump, squirm, or engage in some other activity incompatible with running Examination of behavior on the subsequent trials then showed that animals that had run after receiving the shock made more CRs on the next trial than did animals that had terminated their running or made responses incompatible to running Furthermore, in the avoidance condition he observed that successive avoidances of shock by a conditioned run led to extinction rather than to strengthening of the conditioned running response He therefore interpreted his results as showing that the earlier results of Brogden, Lipman, and Culler were quite consistent with the contiguity principle in learning

The Problem of Secondary Reinforcement

Secondary reinforcement in instrumental conditioning is frequently an uncontrolled variable that reduces comparability of results in instrumental conditioning The Wisconsin laboratory once obtained four Skinner Boxes for rats from a source that will go unnamed When the rats depressed the lever the vending device produced an extremely loud click, and there was no increase in the operant rate of lever pressing Observation of the rats' behavior suggested that they were disturbed by the noise of the vending equipment The equipment was easily altered so as to produce no click at all, but the rats still would not condition A second alteration was made so that the vending device produced a very moderate click, and excellent operant conditioning was then obtained It appears to be generally accepted that the secondary reinforcement provided by audible cues from the vending equipment is of enormous importance in producing operant conditioning in the Skinner Box The noise of the vending mechanism has a better temporal relation to the lever pressing response than does the primary reinforcement associated with obtaining ingesting and digesting the pellet From observations of this sort and many others it is apparent that much of the learning in instrumental situations is dependent upon the nice temporal contingencies possible with secondary reinforcers Troubles arise, however, when secondary reinforcement is present but unsuspected For example, experiments demonstrating great resistance to extinction or pecu

liar temporal relationships between response and reinforcement are usually found to be contaminated by secondary reinforcement or unintended reinforcement of some sort. Because of the ubiquitous character of secondary reinforcement it is virtually impossible to eliminate it from a conditioning experiment.

An excellent illustration of the complications introduced into instrumental learning situations by secondary reinforcement is afforded by Kimble's (1961, p. 150ff) account of the search for the temporal delay of reinforcement gradient. The problem was to investigate the learning of a response as a function of the delay between the response and the reinforcement. This is a valid and fundamental problem and one that seems straightforward enough to promise similar results with different experimental arrangements. Four investigators working on this problem each obtained a smooth, orderly function relating learning to the delay of reinforcement. Their findings are compared in Fig. 9. The functions were

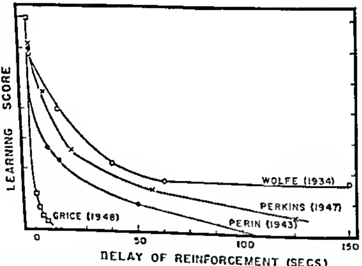

FIG 9. Comparative results of four experiments investigating the effect of delay between response and reinforcement on instrumental conditioning. Wolfe (1934) used a T maze as did Perkins (1947) who tried to eliminate secondary reinforcement from the delay chambers. Grice (1948) used a black white discrimination problem, randomizing correct sides so as to eliminate secondary reinforcement from the proprioceptive stimuli as well as delay box cues Perin (1943) used a Skinner Box, withdrawing the lever during the delay interval

somewhat similar in shape, but gave completely different orders of magnitude for the influence of delay Wolfe (1934) utilized a T maze with a delay chamber in each branch and found that although delays as short as 5 sec. interfered with the learning process, longer delays up to 20 min. continued to produce some learning Perkins (1947) noted that the delay

chambers could provide secondary reinforcement and thus support the learning in spite of long delay Repeating the Wolfe study, but randomly interchanging the delay chambers, he obtained a gradient which fell to a far lower level with the longer delays There remained however, considerable learning with delays as great as 2 min This was attributed by Perkins to the secondary reinforcement provided by the differential proprioceptive cues produced in making the correct or incorrect turn in the T maze Perin (1943) used a Skinner Box in which the response lever was withdrawn after it was depressed and the vending equipment delivered the reinforcement after a specified delay Very little learning was obtained with delays as short as 5 or 10 sec , and extrapolating his curve from 20 sec , Perin predicted that no learning would occur if the delay between response and reinforcement were as great as 40 sec Finally Grice (1948), attempting to eliminate the proprioceptive cues of the Perkins experiment used a black white discrimination problem in which different delays of reinforcement followed the response The positive and negative stimuli were randomly shifted from left to right during the learning trials so that turning left or turning right afforded no constant proprioceptive stimulus related to the correctness or incorrectness of the response Under these circumstances Grice obtained a gradient much steeper than any previously found Delays as short as 2 sec virtually eliminated learning The implications of the discrepancies in results of these four studies are clear Comparability of results cannot be expected until 'irrelevant processes are controlled or eliminated from the experiments It is perhaps of interest to note that in contrast to the variety of results in the foregoing experiments a somewhat analogous function in classical conditioning the effect of the CS-UCS interval has been found to have optimum point at about 0 5 sec in a wide variety of experiments (Kimble 1961, p 156ff)

Summary of the Classification Problem

We are left with this situation When rewards or punishments are used in experiments a complicated mixture of classical and instrumental conditioning takes place And even though the experimenter plans according to one particular paradigm say Pavlovian A conditioning or discriminated omission training which are logically quite different the actual behavior and learning of the *S*s in the two experiments may be very similar indeed being complicated fabrics of classical and instrumental conditioning preparatory and consummatory acts And so the net result is that the protean character of the *S*'s learning dominates the situation so completely that the specific operations of the experimenter are in effect submerged and become of lesser importance It is therefore not surprising that Kimble and also

Konorski (1948, p 243) find tbat the events and phenomena of classical conditioning and instrumental conditioning eventually produce the same sorts of phenomena such as acquisition, extinction, spontaneous recovery, external inhibition, disinhibition, inhibition of delay

There are three ways in which this great similarity of phenomena in widely diversified forms of conditioning may be viewed First of all, the similarity may be more apparent than real, depending upon juxtaposition of isolated examples from the rich variety of experiments available A second view is that the similarities are real and that when learning occurs, regardless of the kinds of reinforcers, the kinds of stimuli, and the kinds of contingencies, the basic process is the same and the behavioral laws are the same And finally a third point of view might be that the issue remains open Pure cases of the paradigm or nearly pure cases either cannot be realized or have rarely been realized so that valid comparisons have not been made This point of view would emphasize that the simple-minded following of the paradigm of a particular classical or instrumental experiment is not sufficient It would require further operational isolation of the stimuli and responses from the usual fabric of behavior to a degree which may, in fact, be impossible in the highly evolved, intact organism Only when the operational paradigms can be realized in actual experiments will comparisons be possible And only then will the issue as to the identity of functional laws of learning in the different conditioning situations be resolved My own view leans toward this last alternative

Conditioning Experiments with Human Subjects

Acknowledging the incompleteness of the classificatory schemes proposed earlier, we shall nevertheless use them as a point of departure for considering the relations of simple conditioning to human learning

Classical Conditioning

Examples of all four types of classical conditioning are readily found in human experimentation These experiments are characterized, as a rule, by fairly rapid learning and phenomena more or less similar to those obtained in the animal experiments Kimble gives a wealth of human conditioning experiments, but it is desirable to cite a few instances of each of the four types as their classification is often proven to be a complicated matter

Pavlovian A conditioning is most obviously exemplified by human salivary conditioning (e g Razran, 1939) In most human conditioning experiments, however, the occurrence of the UCR is not contingent upon *S*'s performance of an instrumental act (such as ingestion of food) so that the

vast majority of the experiments are examples of Pavlovian B conditioning Among these are most human eyelid and GSR conditioning experiments But the human eyelid response can easily be conditioned by avoidance training (Hansche, 1959, Kimble, Mann, & Dufort, 1955, Logan, 1951) and by reward training (Hansche, 1959) as well as by the instructed voluntary method mentioned earlier The human hand withdrawal or finger-withdrawal response is usually conditioned by avoidance training (e g , Hamel, 1919, Schilder, 1929), but can be conditioned by Pavlovian B methods

Owing perhaps to man's tremendous and subtle symbolic capacity crude stimulus substitution of CS for UCS seems to occur but rarely Usually the CR appears to be an "appropriate' response to current stimulation or at least appropriate preparation for what is to come Appearances are often deceiving, however, and the "appropriateness' of human CRs is more likely a product of a remarkably adjustable nervous system than anything more purposive In any event, patellar tendon CRs (Twitmyer, 1902, Wendt, 1930) and conditioned vasomotor responses (Menzies, 1937, Roessler & Brogden, 1943) appear to be instances of crude stimulus substitution as do the organic CRs cited by Bykov (1957) If the pupillary response to light can be conditioned, a debatable point, it might provide another instance of stimulus substitution But most successful pupillary conditioning experiments have involved electric shock as the UCS and hence probably involve mediation of a more complex chain of events than are envisaged in the simple paradigms employed here (Kimble, 1961, p 51)

Instructed voluntary conditioning is, of course, uniquely human Instances, however, occur rarely in conditioning experiments but more commonly as annoying "false reactions" in reaction time and motor skills experiments And, finally, sensory pre conditioning experiments have been carried out with human *S*s (Brogden, 1947, Karn, 1947, Seidel, 1959)

Certain classical conditioning experiments with human *S*s are noteworthy because they are closely related to the more complicated forms of human learning such as concept formation serial and paired associate rote learning, and probability learning Some of these will be briefly mentioned Examples of the more complicated classical conditioning experiments are those of Razran (1939) and Riess (1940, 1946) dealing with semantic conditioning and similar Russian experiments on the second signalling system (Israel Program for Scientific Translations, 1960) where transfer of conditioning was observed to semantically related but not to phonetically related words This resembles concept formation in a primitive way Similarly Wickens (1959) and Shipley (1935) and Dicken (1961) have studied simple seriatim effects connected with mediation and chaining

which resemble serial learning though in a very restricted fashion Also Pavlov (1927, p 117) and others (e g , Rodnick, 1937) have simultaneously conditioned responses to two different stimuli, a primitive analog of paired associate learning And Humphreys (1939) and Grant and Schipper (1952) and others have studied probability learning with classical conditioning methods Even these simple elaborations of the classical conditioning experiments usually produce slow, involved, and somewhat variable learning as compared with the verbal methods to be considered by other participants in the symposium

Instrumental Conditioning

When one searches for human experiments that fit the classificatory scheme for the simple instrumental conditioning experiments described in Table 1, the result is a complete failure Such simple experiments are just not done with normal human *Ss* Human *Ss* undoubtedly learn in these situations, but evidently they learn so rapidly that the experiments can not realistically be brought into the laboratory The experiment is over and done with before the experimenter can observe what happens The experimental problem, posed to an animal, becomes a trivial one when posed to the human *S*

It would appear that in order to obtain sensible, observable learning with human *Ss*, the simple instrumental patterns must be complicated by interposing obstacles to the associative learning Although the arrangement for reinforcing and the nature of the responses may be somewhat similar to the simple experiment, something must be added in order for the problem to require more than one or two trials for human *Ss* to master Some of the ways in which the simple paradigm is complicated are fairly obvious (1) Feeble minded or senescent *Ss* may be used (Bijou & Oblinger, 1960, Bijou & Orlando, 1961, Jeffrey, 1955, Orlando, 1961, Zeaman, House, & Orlando, 1958) (2) The response class may be made subtle or difficult to perform For example, Verplanck (1955) reinforced the statement of 'opinions' during casual conversation so that the *S*, having to emit a class of responses, was essentially engaged in a concept formation experiment Similarly, changes in response class may be made so as to make the response difficult to perform as in the low magnitude responses demanded by Hefferline and his co-workers (Hefferline, Keenan, & Harford, 1959, Hefferline & Keenan, 1961) which were so small as to be unnoticed by the *S* In a sense, this is a motor skill experiment (3) The stimulus may be complicated in such a way that a problem solving or concept formation experiment or probability learning experiment results (4) The number of different S—R connections which must be learned may be multiplied enormously as in paired associate or serial learning,

which brings in an entirely new dimension of interferences such as are rarely found in animal work. (5) Finally, the motivation to learn may be reduced or obscured as in incidental learning and probability learning. All of these procedures result in slower, more observable, learning on the part of the human *S*. The simple associative learning itself, however, is greatly complicated by the obstacles placed in its way in these different kinds of experiments. But, as was noted earlier, the general pattern of stimulating and reinforcing can usually be represented in the classification of Table 1.

Krasner (1958) and Salzinger (1959) have recently reviewed human operant conditioning in relation to verbal behavior. Although the experiments are mostly not impressive, these reviews provide a tremendous number of demonstrations in which various classes of verbal responses are reinforced in various ways, and many of these studies resemble the classes of experiments which will be dealt with later in the symposium. This is not to say that the free operant is easily assimilated into any classification scheme such as those described earlier or that it can be readily compared with verbal learning, concept formation, problem solving, and the like. But an idea of the diversity of these verbal conditioning studies can be obtained from some of the classes of verbal behavior reinforced. These include names of animals, plural nouns, verbs, verbal activity verbs, mother, emotional statements, adverbs or travel verbs, hostile verbs, I and we, pronouns, anti-capital punishment sentiment, various attitudinal responses, acceptance of self verbalization, digit pairs, reported autokinetic effect, Rorschach human responses, movement responses, etc., rate of verbalization, etc. As stated earlier these resemble concept formation experiments.

Comparability of the Forms of Human Learning

If human learning is viewed from the standpoint of the various ways in which the simple instrumental conditioning paradigms are complicated by obstacles put in the path of the human learner, there is an interesting implication for the comparability of the different forms of learning.

Bearing in mind that the learning aspect of all these different experiments presumably involves changes in associative strength of one or more S—R connections, and bearing in mind the foregoing comments on the way human experiments have been complicated beyond the simple paradigms of instrumental conditioning, it is easy to see that there are special difficulties involved in comparison of learning in concept formation, multifold rote learning, problem solving, motor learning, incidental learning and probability learning. For in each instance the associative aspect or

learning aspect has been deliberately embedded in a matrix of other phenomena evolving from the various ways in which the simple instrumental experiment has been complicated so as to give the human *S* a problem worthy of his highly evolved nervous system In most of these experiments the actual S—R association is perhaps the easiest and almost trivial part of the performance The big job for the *S* is to *find* the relevant aspects of the stimulation, the relevant dimension in *concept* formation, to disentangle and retrieve the correct response from all the interfering responses at each moment in *multifold rote learning* to be able to detect relations among a complex of stimuli in *problem solving* to shape a difficult response, evaluating feedback, in *motor learning* to form "correct" S—R associations when there are no guide lines as to what is correct or what is to be learned in *incidental learning* or to form "correct" S—R connections when the feedback is complicated by probabilistic considerations or ambiguities as in *probability learning*

With each of these complications, simple associative learning plays a smaller and smaller role in the total performance of the *S* Therefore, it follows that although all of these forms of behavior are learning, each involves other activities which are quite distinctly different from one experiment to another And insofar as these different forms of behavior constitute major portions of the tasks, then it is to be expected that the behavioral laws derived from the experiments will be quite different, sharing similarities only insofar as they share associative learning which may be only a minor portion of the total task The associative learning may be identical, but the overlay is not

One last word Our chairman, Professor Melton, has raised the point that one might feel, from my terminal remarks, that I regard the kinds of human learning that we study as products of the machinations of psychologists I do not wish to leave this negative impression On the contrary I am completely convinced of the ecological and psychological validity of our experiments in human learning I do believe, however, that in our thinking about learning phenomena we may have adopted in the past a somewhat zoocentric bias as opposed to a more valid anthropocentric approach to learning If so, freeing ourselves of this zoocentric bias may prove helpful The human *S* is capable of more than the animal *S* He learns with great ease things that are difficult or impossible for the animal —and then he goes on to greater complexities of behavior We can present our human *S* with learning problems in conditioning experiments, for example, where he is not particularly efficient and does not far surpass the performance of lower forms, but this is not 'typical' human learning I conclude that a valid psychology of human learning is vastly more complicated than a valid psychology of animal learning, because vastly more numerous, more diversified, and more complicated processes are involved

References

Bijou S W & Oblinger G Responses of normal and retarded children as a function of the experimental situation *Psychol Rep*, 1960 6, 447–454

Bijou, S W, & Orlando R Rapid development of multiple schedule performances with retarded children *J exp anal Behav* 1961 4, 7–16

Brogden W J Sensory pre-conditioning *J exp Psychol* 1939 25, 323–332

Brogden W J Sensory pre-conditioning of human subjects *J exp Psychol* 1947 37, 527–539

Brogden W J, Lipman E A, & Culler E The role of incentive in conditioning and extinction *Amer J Psychol* 1938 51, 109–117

Bykov K M *The cerebral cortex and the internal organs* (Translated and edited by W H Gantt) New York Chemical Publ Co 1957

Cason H The conditioned eyelid reaction *J exp Psychol* 1922 5, 153–196

Cason H Dr Hilgard on the conditioned eyelid reaction *J exp Psychol* 1934 17, 894–899

Dicken C F Connotative meaning as a determinant of stimulus generalization *Psychol Monogr* 1961 75, Whole No 505

Ferster C B & Skinner B F *Schedules of reinforcement* New York Appleton Century-Crofts 1957

Grant D A & Norris E B Eyelid conditioning as influenced by the presence of sensitized Beta responses *J exp Psychol* 1947 37, 423–433

Grant D A & Schipper L M The acquisition and extinction of conditioned eyelid responses as a function of the percentage of fixed ratio random reinforcement *J exp Psychol* 1952 43, 313–320

Grant D A & Schneider D E Intensity of the conditioned stimulus and strength of conditioning I The conditioned eyelid response to light *J exp Psychol* 1948 38, 690–696

Grice G R The relation of secondary reinforcement to delayed reward in visual discrimination learning *J exp Psychol* 1948 38, 1–16

Hamel, I A A study and analysis of the conditioned reflex *Psychol Monogr* 1919 27, Whole No 118

Hansche, W J A comparison of operant, avoidance and classical techniques in the conditioning of the eyelid response Unpublished Ph D dissertation Univer of Wisconsin 1959

Harlow, H F The formation of learning sets *Psychol Rev* 1949 56, 51–65

Hartman, T F, & Grant D A Effects of pattern of reinforcement and verbal information on acquisition extinction and spontaneous recovery of the eyelid CR *J exp Psychol* 1962 63, 217–226

Hartman T F Grant D A & Ross L E An investigation of the latency of instructed voluntary eyelid responses *Psychol Rep* 1960 7, 305–311

Hefferline, R F & Keenan B Amplitude induction gradient of a small human operant in an escape avoidance situation *J exp anal Behav* 1961 4, 41–43

Hefferline, R F Keenan B & Harford R A Escape and avoidance conditioning in human subjects without their observation of the response *Science* 1959 130, 1338–1339

Hilgard E R Conditioned eyelid reactions to a light stimulus based on the reflex wink to sound *Psychol Monogr* 1931 41, Whole No 184

Hilgard E R & Marquis D G Conditioning and learning New York Appleton Century 1940

Humphreys L G The effect of random alternation of reinforcement on the acquisition and extinction of conditioned eyelid reactions *J exp Psychol* 1939 25, 141–158

Israel Program for Scientific Translations *The central nervous system and*

behavior Selected translations from the Russian medical literature Bethesda, Maryland National Institutes of Health, 1960

IVANOV SMOLENSKY, A G Metodika issledovaniya uslovnykh reflcksov u cheloveka (Methods of investigation of conditioned reflexes in man) Moscow Medgiz (Medical State Press), 1933

JEFFREY, W E New technique for motivating and reinforcing children *Science,* 1955, 121, 371

KARN, H W Sensory preconditioning and incidental learning in human subjects *J exp Psychod*, 1947, 37, 540–544

KIMBLE, G A., MANN, L I, & DUFORT, R H Classical and instrumental eyelid conditioning *J exp Psychol*, 1955, 49, 407–417

KIMBLE, G A Hilgard and Marquis' Conditioning and learning New York Appleton Century-Crofts, 1961

KONORSKI, J Conditioned reflexes and neuron organization Cambridge, England Cambridge Univer Press 1948

KRASNER L. Studies of the conditioning of verbal behavior *Psychol Bull*, 1958, 55, 148–170

LEVY, C M, CLARK, A, & GRANT, D A Classically conditioned discrimination reversal *J exp Psychol*, 1963, 66, (In press)

LOGAN, F A A comparison of avoidance and non avoidance eyelid conditioning *J exp Psychol* 1951, 42, 390–393

MARQUIS, D G, & PORTER, J M, JR. Differential characteristics of conditioned eyelid responses established by reflex and voluntary reinforcement *J exp Psychol*, 1939 24, 347–365

MCCLEARY, R. A. Response specificity in the behavioral effects of limbic system lesions in the cat *J comp physiol Psychol*, 1961, 54, 605–613

MENZIES P Conditioned vasomotor responses in human subjects *J Psychol*, 1937, 4, 75–120

MILLER, N E Learnable drives and rewards In S S Stevens (Ed) *Handbook of experimental psychology* New York Wiley, 1951

ORLANDO, R. Component behaviors in free operant temporal discrimination *Amer J ment Defic*, 1961, 65, 615–619

PAVLOV, I P *Conditioned reflexes* (Translated by G V Anrep) London Oxford Univer Press, 1927, reprinted New York Dover, 1960

PERIN, C T A quantitative investigation of the delay of reinforcement gradient *J exp Psychol*, 1943, 32, 37–51

PERKINS, C. C., JR. The relation of secondary reward to gradients of reinforcement *J exp Psychol*, 1947, 37, 377–392

RAZRAN G A quantitative study of meaning by a conditioned salivary technique (semantic conditioning) *Science* 1939 90, 89–91

RAZRAN, G The observable unconscious and the inferable conscious in current Soviet psycholoy Interoceptive conditioning semantic conditioning and the orienting reflex. *Psychol Rev* 1961 68, 81–147 (a)

RAZRAN, G Recent Soviet phyletic comparisons of classical and of operant conditioning *J comp physiol Psychol* 1961, 54, 357–365 (b)

RIESS, B F Semantic conditioning involving the galvanic skin reflex *J exp Psychol*, 1940 26, 238–240

RIESS, B F Genetic changes in semantic conditioning *J exp Psychol*, 1946, 36, 143–152.

RODNICK, E. H Does the interval of delay of conditioned responses possess inhibitory properties? *J exp Psychol* 1937 20, 507–527

ROESSLER, R L., & BROGDEN, W J Conditioned differentiation of vasoconstriction to subvocal stimuli *Amer J Psychol* 1943, 56, 78–86

SALZINGER, K Experimental manipulation of verbal behavior a review *J gen Psychol*, 1959, 61, 65–94

SCHILDER P Conditioned reflexes *Arch Neurol Psychiat* 1929, 22, 425–443

SEIDEL, R. J A review of sensory preconditioning *Psychol Bull*, 1959, 56, 58–73

SHEARN, D Does the heart learn? *Psychol Bull* 1961, 58, 452–458

SHEFFIELD, F D Avoidance training and the contiguity principle *J comp physiol Psychol*, 1948, 41, 165–177

SHEFFIELD, F D, & TEMMER, H W Relative resistance to extinction of escape training and avoidance training *J exp Psychol* 1950 40, 287–298

SHIPLEY, W C. Indirect conditioning *J gen Psychol* 1935 12, 337–357

SKINNER, B F Two types of conditioned reflex A reply to Konorski and Miller *J gen Psychol*, 1937, 16, 272–279

SKINNER, B F The behavior of organisms An experimental analysis New York Appleton Century, 1938

SKINNER, B F A case history in scientific method In S Koch (Ed), *Psychology A study of a science* New York McGraw Hill 1959 Vol 2, pp 359–379

SPENCE, K W, A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations *Amer Psychologist* 1958, 13, 131–141

SPENCE, K W, & ROSS, L. E. A methodological study of the form and latency of eyelid responses in conditioning *J exp Psychol* 1959, 58, 376–385

SPENCE, K W, & TAYLOR J Anxiety and strength of the UCS as determiners of the amount of eyelid conditioning *J exp Psychol* 1951, 42, 183–188

THORNDIKE E L. The law of effect. *Psychol Rev*, 1933 40, 434–439

TWITMYER, E B A study of the knee jerk Ph D thesis, Univer of Pennsylvania Privately printed Philadelphia Winston 1902

VERPLANCK, W S The control of the content of conversation Reinforcement of statements of opinion *J abnorm soc Psychol*, 1955, 51, 668–676

WATSON, J B, & RAYNER, R Conditioned emotional reactions *J exp Psychol*, 1920 3, 1–14

WEBER, H, & WENDT, G R Conditioning of eyelid closures with various conditions of reinforcement *J exp Psychol* 1942, 30, 114–124

WENDT, G R An analytical study of the conditioned kneejerk *Arch Psychol*, N Y, 1930 19, Whole No 123

WESTCOTT, M R, & HUTTENLOCHER, J Cardiac conditioning The effects and implications of controlled and uncontrolled respiration *J exp Psychol* 1961, 61, 353–359

WICKENS, D D Conditioning to complex stimuli *Amer Psychologist* 1959, 14, 180–188

WOODWORTH R S, & SCHLOSBERG H *Experimental psychology* (Rev ed) New York Holt, 1954

WOLFE, J B The effect of delayed reward upon learning in the white rat *J comp physiol Psychol* 1934 17, 1–21

WYCKOFF, L. B, JR The role of observing responses in discrimination learning Part I *Psychol Rev*, 1952, 59, 431–442

ZEAMAN, D, HOUSE, B J, & ORLANDO, R Use of special training conditions in visual discrimination learning with imbeciles *Amer J ment Defic* 1958, 63, 453–459

ZENER, K The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response *Amer J Psychol*, 1937, 50, 384–403

Categories of Learning and the Problem of Definition.

COMMENTS ON PROFESSOR GRANT'S PAPER

GREGORY A KIMBLE[1]
Duke University

Stimulated by Professor Grant's excellent discussion of classical and instrumental conditioning, I have been led to view the problem of establishing categories of learning more generally than I have in the past, and am glad to have this opportunity to record some resulting ideas here The first part of what I have to say is of a general methodological nature

Let us begin by recognizing that the problem of establishing a taxonomy for learning is a definitional problem in which we must deal with the same issues as are involved in any other such problem in psychology In very brief It is essential that whatever distinctions we draw rest on sound operational grounds and that these operationally drawn distinctions have some relevance to behavior (Kimble, 1953) In what may be opposition to Professor Grant's point of view, I would hold that the operational identification of categories of learning is not a mere alternative to other conceivable procedures for the establishment of such categories, to be resorted to when other modes of definition fail I would hold, instead that such operational definition is, in a sense, separate from the alternative possibilities, that, speaking generally, such definition provides for an initial identification of concepts (and distinctions among concepts), that, speaking more specifically to the point of our present discussion, the operational specification of types of learning parallels the operational definition of learning itself, and that the other distinctions, to which Professor Grant refers, are like the dependent variables associated with any concept, serving if dependable to validate the original operational specification This last series of points provides a sort of skeleton outline of the content of this paper

Before turning to the problem of identifying varieties of learning, it will be useful to deal briefly with the more general concept, that is, with learning itself Elsewhere (Kimble 1961) I have defined learning as a relatively permanent change in a behavioral potentiality which occurs as a result of

[1] The preparation of this paper was facilitated by a grant, NSF-G 7079 from the National Science Foundation.

reinforced practice Put into a familiar form, this definition may be schematized as in Fig 1 with "reinforced practice" on the independent variable side of the diagram and with "a relatively permanent change in behavior" appearing on the other, dependent variable, side of the diagram The concept, learning, occupies the typical intermediate position of an *intervening variable*

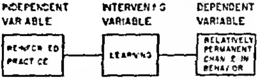

Fig 1 The status of the concept learning as an intervening variable

As I have also spelled out elsewhere (Kimble, 1956, 1961), this definition does a fairly satisfactory job of distinguishing learning from such processes as adaptation, fatigue, experimental extinction, changes in motivation, and maturation which also lead to changes in behavior The term reinforcement which occurs in this definition has no particular theoretical meaning (Kimble, 1961, pp 5–6) It refers, operationally, simply to any state of affairs which, introduced appropriately into a situation, leads to the strengthening of some S—R relationship

The point which I want to make next is that distinctions among different kinds of learning may profitably be viewed in terms of this same paradigm It is necessary that whatever distinctions we decide to make be based on firm operational grounds, and that the operationally defined distinctions have some relevance to behavior Failing to meet the former criterion makes it impossible to distinguish among kinds of learning reliably, failing to meet the latter renders whatever operational distinctions we make insignificant

The specific problem which Professor Grant discusses in his paper is that of the distinction between classical and instrumental conditioning Approaching this problem from the point of view just developed leads to the construction of a second diagram, which is an elaboration of the one presented in Fig 1 This diagram, which appears in Fig 2, is an attempt to relate the two kinds of conditioning to the more general concept learning, and to raise in summary form some of the questions to which we shall want, later on, to address ourselves regarding the differences between the two forms of learning As the left hand side of that diagram shows, it is possible to distinguish between classical and instrumental conditioning on operational grounds, in terms of the contingency, or lack of it between a response and a reinforcer The materials on the right hand side of the diagram raise several questions concerning the significance for the student

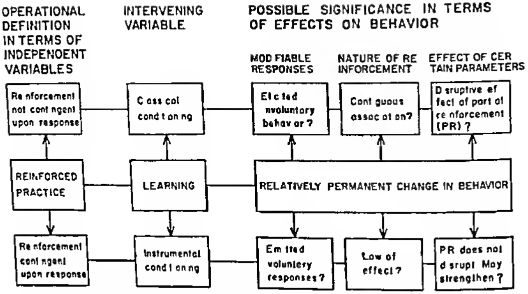

FIG 2 Diagram to illustrate how the distinction between classical and instrumental conditioning entails the same considerations as the definition of learning

of learning, of the operational distinction between classical and instrumental conditioning The general idea which this elaborated diagram attempts to convey is that classical and instrumental conditioning both qualify as examples of learning, but that they *may* differ (this is an empirical question) in several ways, including the kinds of responses which enter into classically and instrumentally conditioned associations, the nature of reinforcement as it applies to the two forms of learning, and the operation of certain variables which may influence the two processes differently The majority of the remainder of this paper will be devoted to a more detailed treatment of the items which appear on the right hand side of Fig 2, in an effort to determine the validity of the distinction between classical and instrumental conditioning

CAN THE SAME RESPONSES BE CONDITIONED BOTH CLASSICALLY AND INSTRUMENTALLY?

In my revision of *Hilgard and Marquis Conditioning and Learning* (1961) I suggested that the operational distinction between classical and instrumental conditioning may correspond to a behavioral dichotomy according to which some responses are conditionable only by the methods of classical conditioning and others are conditionable only by the methods of instrumental conditioning This suggestion is repeated as the first item on the right hand side of Fig 2 In a moment I should like to offer a brief defense of this possible distinction Before doing so, however, perhaps I

should mention that I recognize that this is an extreme and even implausible position Doubts as to its validity arise, if nowhere else, from the general observation that Nature appears to abhor dichotomies much as she abhors vacuums On the other hand, and in spite of such doubts, I know of no firm evidence to refute the position Perhaps you will not think it too odd that I can simultaneously doubt the validity of a point of view and still feel that it stands a better chance of being right than any competing position

Also, as a sort of preliminary to the main point I should like specifically to dissociate myself from Mowrer s (1960) position Mowrer, of course, makes the same distinction as that which I have proposed He however, makes the further assumption that classically conditionable responses are always autonomic responses and that instrumentally conditionable responses are always skeletal The classical conditionability of such skeletal reflexes as the eyeblink seems to me an adequate refutation of one half of this hypothesis Whether, conversely, autonomic responses can be conditioned instrumentally is one of the fascinating issues in learning today I personally think they cannot be, but this is an active area of research and the evidence is not in More of this in a moment The point here is just that I doubt that the distinction drawn by Mowrer is exactly right

Since the publication of *Conditioning and Learning* in which I suggested the distinction outlined above, I have received several thoughtful communications objecting to it Although each of the objections is a reiteration of an old argument, there seems to be real point in a brief reconsideration of one or two of the more persistent The first of these objections has been that, in some cases, it seems possible that the same response may be modified both classically and instrumentally This would constitute factual evidence against the theory that classically conditionable reactions cannot be conditioned instrumentally, and *vice versa* My reaction to this criticism is simply that, so far, I have not been convinced that it is correct On the one hand, there is as yet no convincing evidence that responses which are normally involuntary, such as heart rate changes and the GSR can be directly conditioned instrumentally, although they can easily be modified classically In cases where the GSR or heart rate have allegedly been conditioned instrumentally, it is always very easy to detect possible skeletal mediating reactions which constituted the true response in question A little tenseness in the case of the GSR, or a catch of breath in the case of the heart rate, is all that would be required A really telling demonstration of the instrumental conditionability of such responses might consist of the development of such instrumental CRs in curarized animals But so far, to the best of my knowledge no such experiment has been done

Another example, that of the conditioned eyeblink which also super

ficially appears to be modifiable in both ways (Moore & Gormezano, 1961), requires a different explanation In this case, the classically and instrumentally conditionable eyeblinks, in my opinion, are different responses More specifically, I believe that the instrumentally conditioned eyeblink is that which investigators in this field call the voluntary blink and that the classically conditionable reaction is of a different sort Figure 3 presents a comparison of typical voluntary and involuntary responses

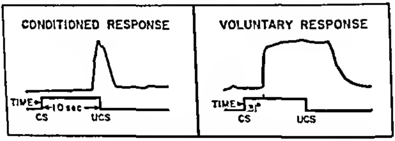

Fig 3 The characteristic forms of involuntary (left hand tracing) and voluntary (right hand tracing) anticipatory eyeblinks

which occur in the eyelid conditioning situation in the form of tracings of sample records The distinguishing features of the voluntary response are its short latency, rapid recruitment, and protracted duration In these characteristics it is quite different from the reaction which is the usual object of investigation in the eyelid conditioning experiment The differences apparent in the tracings in Fig 3 have contributed to the conclusion reached by the majority of investigators in eyelid conditioning that these two responses are of basically different types which probably obey different laws Support for this conclusion comes from two lines of evidence which have not been described sufficiently often in the literature, probably because they are matters of common laboratory knowledge rather than experiment, but, in general (1) When *Ss* are instructed to blink voluntarily to the CS, the response usually takes the voluntary form, similarly, when *Ss* who give many responses of this form are questioned, they often report having blinked voluntarily (2) When *Ss* are instructed not to blink voluntarily, or to "let their blinking take care of itself," responses are less often of the voluntary form, similarly, when *Ss* who show the nonvoluntary form of response are questioned, they seldom report having blinked "on purpose" and sometimes are surprised to discover that they have blinked in anticipation of the UCS at all So far, differences in the principles which may govern these two kinds of behavior have not been adequately investigated Informal evidence and some experimentation (Kimble, 1961, p 101) suggest that the voluntary response develops full strength much

more rapidly than the involuntary variety, and may attain a level of 100% occurrence after a few trials

A second criticism of the hypothesis of the independence of classically conditionable and instrumentally conditionable responses is most fully developed by Hebb (1956) in a paper of which I was unaware until very recently Briefly, what Hebb does is to emphasize the rewarding consequences of the responses to many classical unconditioned stimuli and to maintain that certain forms of conditioning usually regarded as classical are actually instrumental Salivation is the best example This is because (as others have also recognized) salivation no doubt increases the palatability of food (or dilutes acid when that is the UCS) and constitutes a rewarding state of affairs which cannot be separated from any other effects which the unconditioned stimulus may have—except conceptually What I would want to emphasize here is the importance of achieving this conceptual separation

One of the major pitfalls which psychology must avoid derives from what Stuart Chase (1938) long ago called "the tyranny of words" In psychological thinking, what is involved is an overly common naive faith in the exact correspondence of language and reality To a very large extent, we seem to operate in terms of the formula one word, one item of reality In the present context, this leads to an oversimplified conception of classical conditioning, in which it is assumed that a single event (abstractly referred to as the UCS) influences a single aspect of behavior In my opinion, both assumptions in this simple view are wrong For most reinforcers are very complex events and the behavior which they modify is also complex In the case of salivary conditioning, the situation is something like this The food or acid used as a reinforcer does at least two things (1) It elicits salivation and, no doubt, in addition, a variety of emotional reactions (2) Indirectly, through the elicitation of salivation, it also increases the palatability of food or dilutes the acid and reduces its unpleasant sour taste I believe that it is very nearly certain that these two different effects of food have different influences upon behavior, and modify the organism in very different ways The first process (the elicitation of salivation and its emotional accompaniments) I believe to be responsible for the establishment of the classically conditioned salivary reflex The second effect of the reinforcer (its effect upon the taste of food) has different consequences This latter influence is not upon salivation however, but upon the host of other reactions which the experimental animal tends to develop More concretely, the obviously skeletal reactions of looking toward the place where food appears, or approaching it in the free salivary conditioning situation employed by Zener (1937) and men

tioned by Grant, in my opinion, are instrumentally conditioned and depend upon the second (essentially rewarding) effect of the food I suspect that it would not be difficult at all to show that these reactions are very different depending upon whether the classical CR they accompany is established with food or acid as the UCS Such a demonstration, again, would argue for a treatment of salivation and the instrumental accompaniments of salivation as very different kinds of learning which happen to occur together in the most classical of classical conditioning situations

Subclasses of Classical Conditioning

Although I intend to return to the major distinction between classical and instrumental conditioning, it will be convenient, with the background which has now been developed, to make a series of comments upon the varieties of classical conditioning proposed by Professor Grant If I had been doing it, I would not have arrived at the same categories as Professor Grant did, but this hinges in an important way upon the previous commitment just described to a distinction between classical and instrumental conditioning in terms of the responses which can be conditioned with these procedures

First, however, to remind you of the set of distinctions which Professor Grant makes, they are (1) Pavlovian A (that is, ordinary salivary) conditioning (2) Pavlovian B, conditioning typified by nausea elicited by the preparations for morphine injection, (3) anticipatory instructed conditioning and (4) sensory preconditioning Such objections to these categories as I have involve the last two varieties Pavlovian A conditioning I would classify in my own scheme as conditioning employing a positive reinforcer (the nature of positive reinforcement being independently defined in a Thorndikean way as proposed by Grant) Pavlovian B is a complex and not very good example of classical conditioning employing a negative reinforcer To these, I would consider adding the two forms of classical conditioning proposed by Mowrer (1960) in which the CS signals the absence or impending termination of a positive reinforcer in one case and of a negative reinforcer in the other Whether these forms of conditioning actually occur, of course, is a moot question But one can suggest them on purely operational grounds, because it is perfectly obvious that the physical arrangements implied in each procedure are possible to establish If nothing else, this analysis suggests lines of interesting experimentation It also seems to fit very well into Professor Grant's classification of instrumental conditioning and might be recommended to him on this ground

Personally, I would prefer to deal with the other two varieties of "classical" conditioning proposed by Professor Grant in the following way The anticipatory instructed variety, in my opinion, is actually instrumental con-

ditioning The reasons for thinking of such conditioning as instrumental again reflect my hypothesis about the separability of responses which can be conditioned classically and instrumentally Those involved in anticipatory instructed conditioning are voluntary (by definition) which, in my opinion, means that they can be conditioned only instrumentally and not classically

Sensory preconditioning I would consider to be a simple case of classical conditioning in which the two CSs actually bear to each other the relationship of CS and UCS On this basis it is possible to view the preconditioning phase of the experiment as an instance of true conditioning The usual experiment on sensory preconditioning however, makes no provision for recording the responses involved There are many implications of this idea, few of which have been explored experimentally In general however, adopting this point of view mediates the prediction of a number of effects of sensory conditioning It should, for example, be possible to show that there actually are responses to at least the stronger of the two stimuli used in the sensory preconditioning experiment and that sensory preconditioning does not occur unless there are such responses In cases where the two stimuli do differ in strength sensory preconditioning should be possible only when the stronger is presented second in original training otherwise the experiment would be one in backward conditioning The interval separating these two stimuli should be important and it should be possible to demonstrate extinction of the sensory preconditioned response To the extent that there is evidence on these points (e g Coppock, 1958), that evidence tends to be positive

Classical and Instrumental Conditioning and the Nature of Reinforcement

The second possible difference between classical and instrumental conditioning suggested in Fig 2 involves the major question on which research has been done in this area, that is, on the question of the nature of reinforcement There will not be space here for me to review the work which has been done in this area Moreover, there seems to be no real need to do so, since discussion of this point is available in a number of sources (e g, Kimble, 1961, Mowrer, 1960) For the sake of completeness however, it may be worth mentioning that two issues are at stake The first of these involves a decision on the part of the theorist whether to adopt a uniprocess or multiprocess theory of learning The uniprocess views need not concern us here because, by definition they do not recognize a distinction between classical and instrumental conditioning except on operational grounds This as we have already seen is tantamount to making no

distinction at all As Fig 2 suggests, I am inclined to think (and I believe that Professor Grant agrees) that there are important differences between the two forms of learning, one of which is under discussion

Having accepted a two-process position, the question to which we must turn next concerns possible differences in the nature of reinforcement for the two varieties of learning. Here the most popular choice, and that toward which I incline, is to view classical conditioning as learning which results from a process of contiguous association, and to view instrumental learning as something which obeys the law of effect There are, of course, various opinions as to the nature of the mechanism underlying the operation of the law of effect The evidence does not warrant a choice among the various possibilities, and this is not the place to pursue the question further Here our only immediate interest is in the question of the applicability of the law of effect to classical conditioning The two-process theory in its most typical version holds that the law of effect does not apply to classical conditioning

The most important line of evidence which appears to support this point of view is that experiments dealing with the effect upon conditioning of the duration of noxious unconditioned stimuli have shown that this variable has little or no influence upon the level of conditioning The significance of this is that, with noxious reinforcers, the only way in which reward (involved by definition in the law of effect) could operate to strengthen a response followed by such a stimulus would be if the important event were the cessation of the stimulus If this were so, the use of a long-duration UCS would be the same as postponing the occurrence of reinforcement On other grounds, we know that delaying reinforcement lowers the level of conditioning This leads to the prediction that classical conditioning should be inferior with a protracted UCS by comparison with the level of conditioning obtainable with a short UCS Since this effect does not occur in classical conditioning (e g , Wegner & Zeaman, 1958), the suggestion is that the theory from which the prediction derives is wrong More specifically, reinforcement is not provided by the cessation of the noxious stimulus and, therefore, this is not an instance in which the effect strengthens the response in question

Are Laws of Classical Conditioning and Instrumental Conditioning the Same?

In discussing this problem Professor Grant offers three alternatives and apparently leans slightly, himself, to the view that the question cannot be answered at the present time I agree, but perhaps for somewhat different

reasons The central points which I have in mind are these In the first place, there seems little reason to doubt that experiments in instrumental and classical conditioning have yielded an impressive array of similar phenomena Such features of learning as acquisition, extinction, spontaneous recovery, stimulus generalization, response generalization, and discrimination are shared by both This could easily lead to the conclusion that, at least at one level, the laws of both kinds of learning must be the same Probably the majority of experts in the field of learning would accept such a conclusion Against its unquestioned acceptance, however, there stands this fact In setting up an experiment in instrumental conditioning, it is impossible to avoid the inclusion of classical conditioning This is because to insure the occurrence of instrumental conditioning, it is necessary to provide for reinforcement The rat learning to press the bar in the Skinner Box must be rewarded with food, water, avoidance of shock, or something else Moreover, it is perfectly obvious that this reinforcement must occur in some situation In the Skinner Box example, food is administered in the compartment This means that the stimuli in the box are paired with reinforcement after the manner of classical conditioning It seems likely, therefore, that classical conditioning occurs as a terminal element[3] on each trial or with each series of free responses in the Skinner Box situation Given that the various phenomena just mentioned are characteristic of classical conditioning, it is to be expected that they would occur in the case of classically conditioned responses which contaminate instances of instrumental learning Moreover, there is a possibility that the entire explanation for these phenomena in instrumental learning reduces to their occurrence in the classically conditioned element which terminates the trial That is, it is at least conceivable that such phenomena as extinction, stimulus generalization, and the like occur in instrumental learning experiments only because of the occurrence of these processes in the inevitable classical conditioning component of the experiment Whether this is so or not depends, as Professor Grant suggests upon the results obtained

[3] In editing this paper, Professor Melton raised the question of whether the classical conditioning component must be limited to the *terminal* element of the chain of instrumental acts It seems to me that classical conditioning must involve as the UCS an element of some emotional significance Thus in most instrumental reward situations the associations formed by classical conditioning would involve the food as a UCS and other aspects of the experimental situation as CSs The strength of such conditioning would be greatest for neutral stimuli which occur briefly in time before the food This would effectively limit the process to the terminal components of the response sequence On the other hand as we shall see such conditioning might *generalize* to other portions of the situation and influence behavior in areas distant from the goal

when instrumental learning is analyzed into its component processes, and the laws of both forms of learning have been independently studied So far, not much along these lines has been done

At the present time, the chief variable which seems to provide a possible entering wedge into this problem is that of the schedule of reinforcement This possibility was suggested as the final item on the right-hand side of Fig 2 From Pavlov on, there has been evidence that reinforcing less than every trial is a peculiarly destructive procedure in classical conditioning Pavlov (1927), for example, found that reinforcing on less than a one in three schedule made conditioning of the salivary response impossible In instrumental conditioning by contrast, performance frequently is weakened very little by the use of much smaller ratios of reinforcement In fact, the evidence suggests that, when a sufficient number of trials is examined, schedules of partial reinforcement may facilitate rather than interfere with performance in the instrumental learning situation (Goodrich, 1959, Weinstock, 1958)

Experiment I[3] in the paper by Goodrich is a study whose implications I find very interesting Goodrich ran rats in a straight runway on 100% and 50% schedules of reinforcement and obtained three separate measures of running speed, in the starting box, main alley, and goal box The results of the experiment showed that (1) Early in learning, running speed was slightly faster with the 100% schedule than with the 50% schedule for all regions of the alley (2) After about 15 and 25 trials, for starting speed and alley speed, respectively, the 100% and 50% curves crossed, with the partial groups displaying superior performance (3) This crossing did not occur for speed of running in the goal box, the 100% group remained superior

Although Goodrich does not offer such an explanation, these results may reflect the differential effect of partial schedules of reinforcement upon classical and instrumental conditioning for which I have just been arguing The ultimate superiority of performance under the partial schedule for the first two segments of the maze obviously shows that instrumental learning suffers no decrement under such schedules and may be superior The fact that running speed remains superior under the continuous schedule in the goal box may reflect a superior conditioning of the classical component to cues in the goal box under continuous reinforcement However, the

[3] Experiment II was complicated by an alteration in apparatus which may have eliminated the difference in behavior of continuous and partial groups in the region of the goal box for reasons that are more physical than psychological The fact that the results of Friedes (1957) are like those of Goodrich s Experiment I supports this interpretation.

latter point, which is crucial, raises a difficult point of interpretation, because the response being measured is clearly an instrumental response and not a classical one. This problem may be resolved when we know more about the interaction between the classical and instrumental components of the behavior occurring in a single situation.

The results of the Goodrich experiment itself, suggest one possible form of such an interaction. Let us assume with Hull (1952) that the classical component energizes instrumental behavior through a contribution to the animal's motivation (incentive K) and (going beyond Hull) that this incentive is made effective by generalization from the goal situation to other places in the apparatus. From Goodrich's results it seems possible that, initially, this generalization is quite wide, extending even to the starting box in the alley and producing superior performance under a continuous schedule, because of superior classical conditioning under such schedules. Increasing practice, however, leads to a gradual restriction of the range of generalization so that the superiority of performance under the continuous schedule disappears, first for performance in the starting box and then for performance in the alley, remaining only in the goal box where the cues to the classical CR are most strongly conditioned. Obviously, an analysis of this kind is *ad hoc* and entails a number of unproven assumptions (such as that the range of generalization narrows with practice when the dimension of generalization is of the spatio temporal variety). It also begs the question (since it assumes the superiority of classical conditioning under continuous reinforcement which is the point to be demonstrated). On the other hand, the results of this experiment may be taken as providing indirect support for the present interpretation and also as illustrating possible applications.

More direct evidence on the effect of partial schedules of reinforcement upon classical conditioning is undesirably sparse. Such evidence as there is (e g , Moore & Gormezano, 1961) indicates that partial reinforcement interferes quite seriously with the conditioning of the eyeblink, the response upon which most of the work on the question has been done. We badly need comparisons of continuous and partial reinforcement which employ other classically conditionable responses as well as attempts to demonstrate the differential influence of other variables upon classical and instrumental conditioning. Such data would contribute importantly to reaching a decision on the question of whether classical and instrumental conditioning are actually two separate forms of learning. As I have tried to indicate in this paper, my belief is that they are indeed different but the evidence does not really warrant a completely confident statement to that effect.

Summary

The problem of establishing a satisfactory taxonomy for learning is a definitional problem involving the same issues as other problems of this type (1) There must be a clear operational distinction among categories of learning and (2) these operationally established categories must have some significance for behavior

In the case of the distinction between instrumental and classical conditioning the first of these requirements is clearly satisfied in terms of the contingency or lack of it between the response and reinforcement The second requirement poses certain issues which have been of central interest to the psychology of learning for several decades Possibly significant differences between classical and instrumental conditioning are these (1) It may be that classical conditioning is a process which applies only to involuntary elicited behavior whereas instrumental conditioning applies to voluntary emitted behavior (2) The evidence suggests that classical conditioning involves the establishment of associations as a result of S—S or S—R contiguity and that instrumental conditioning obeys the law of effect (3) It seems likely that partial reinforcement is a destructive procedure with classical conditioning and not with instrumental conditioning

This last item suggests that the laws of classical and instrumental conditioning may not be exactly the same This proposition is difficult to assert with confidence however because of the great difficulties of separating the two forms of learning in the actual experimental situation

References

Chase S *The tyranny of words* New York Harcourt, Brace and World 1938

Coppock, W J Pre-extinction in sensory preconditioning *J exp Psychol* 1958 55 213 219

Fitzdes D Goal box cues and pattern of reinforcement. *J exp Psychol* 1957 53 361 372.

Goodrich K P Performance in different segments of an instrumental response chain as a function of reinforcement schedule *J exp Psychol* 1959 57, 57–63

Hebb D O The distinction between "classical" and "instrumental" *Canad J Psychol* 1956 10 165–166

Hull C L. *A behavior system* New Haven Yale Univer Press 1952

Kimble G A Psychology as a science *Scientific Monthly* 1953 77 156–160

Kimble G A *Principles of general psychology* New York Ronald Press 1956

Kimble G A *Hilgard and Marquis' conditioning and learning* New York Appleton-Century-Crofts 1961

Moore J W & Gormezano I Yoked comparisons of instrumental and classical eyelid conditioning. *J exp Psychol* 1961 62, 552 559

Mowrer O H *Learning theory and behavior* New York Wiley 1960

PAVLOV, I P *Conditioned reflexes* (Translated by G V Anrep) London Oxford Univer Press, 1927
WAGNER, NORMA, A ZEAMAN, D Strength of cardiac CR s with varying unconditioned stimulus durations *Psychol Rev*, 1958, 65, 238–241
WEINSTOCK, S Acquisition and extinction of a partially reinforced running response at a 24 hour intertrial interval *J exp Psychol*, 1958, 56, 151–158
ZENER, K The significance of behavior accompanying conditioned salivary secretion for theories of the conditioned response *Amer J Psychol*, 1937, 50, 384–403

The Representativeness of Rote Verbal Learning

BENTON J UNDERWOOD
Northwestern University

A poll taken among behavioral scientists would probably show mild firmness of belief that there is some continuity in the laws of behavior across all phyla in the animal kingdom Were we to inquire about a family within a phylum a somewhat firmer belief in the continuity would be manifest If we go to a given genus the belief becomes still more firm, and within the species, *Homo sapiens*, it would be firm indeed If we follow this increasingly restrictive path one step further we arrive at one of the purposes of this conference, namely, to analyze, and eventually to make explicit how commonality of behavior is exhibited when we study the *learning* of the human organism in different situations Assuming a positive correlation between firmness of belief and fact, it would seem that the conclusion is inescapable that there are laws of human learning which at once supersede but include the learning exhibited in any particular situation These laws, it would seem, *must* be present and even a cursory analysis would show what they are Chapter headings, such as Conditioning, Rote Learning Motor Learning, and so on, we would say, are just that and no more, they should not be taken to indicate that we are dealing with human learnings But, on second thought, perhaps the commonality is not as apparent as it appears that it should be Perhaps if it were apparent this conference would not be required Still, some one has to catalogue the commonality, some one has to point out the basic laws, the common dimensions, the common phenomena, or the identities of the learnings appearing under different chapter headings Surely a careful consideration will divulge the continuity that is said to exist

The assumption of continuity does not, of course, demand that we find evidence for complete or perfect continuity wherever we look The manifestation of the laws may be different in different situations, certain research paradigms may not allow certain laws to be operative at all, complex interactions may mask the effects of basic laws And that may be the situation we face The inevitable pyramiding of research has heightened

the fragmentation in the study of human learning, and it may not be easy to bring together the fundamental factors involved in all forms of human learning and thus bring concordance between belief and fact

Approaches to a Taxonomy of Human Learning

Some Possible Approaches

There are many approaches which might be used to express the relationships among research findings for all forms of human learning Undoubtedly the most elegant way would be in terms of theory A general theory of human learning in which the particular findings in each area are shown to be deductions from the master set of statements and relationships is clearly an ideal solution No such system is available Since the scientist is never content with particulars we do have many, many attempts to apply theoretical notions across areas for limited kinds of phenomena All of the contemporary models (the favored term today) are presumed to have greater generality than the particular situation in which they have been employed, and it is quite possible that some of these will eventually provide the basis for a general theory of human learning

Another way of attempting to express the continuity for all human learning is in terms of phenomena produced by comparable operations Thus, can the operations defining extinction in eyelid conditioning be duplicated in verbal learning, in motor learning, in concept formation, and so on, and, if so, do the same phenomena result from these operations? This approach simply asks whether manipulable variables affect learning in the same way, regardless of the task involved Systematic work of this nature across many variables would be of inestimable value in constructing a general theory of human learning In fact, this approach has been rather widely used, although not in any systematic sense and not necessarily with the intent of unifying areas The approach has certain limitations and certain elements of danger The limitations come from the fact that there are some tasks which are simply not of such a nature that certain variables can be introduced For example, it would seem difficult to manipulate meaningfulness on a pursuit rotor in the same sense that this variable is manipulated in verbal learning Or, what operations in problem solving are comparable to variations in intensity of the conditioned stimulus in classical conditioning? Of course we have to say that if a given variable in one situation has no possible counterpart in another then it simply is not a relevant variable in the latter However, if this asymmetry is widespread among tasks any systematic program to manipulate variables across tasks becomes somewhat less than systematic

The dangers of the method are two in number We have no criteria for

clearly specifying when operations are comparable when tasks differ In deed, some might say that if the tasks differ the operations cannot be comparable The complete entertainment of such counsel would remove all hope of assessing empirically the continuity of human learning across tasks, and we should not subscribe to it But even with less rigid criteria of saying when operations are and are not comparable, there is still an issue of concern For example, in classical conditioning there is a body of literature on the effects of varying percentage of reinforcement, i e the proportion of the trials on which the unconditioned stimulus is presented Some studies (e g , Schulz & Runquist, 1960) have been done with paired associate learning in which the proportion of trials on which the correct response was shown was varied The results differ somewhat from those obtained with classical conditioning Is this because the operations are not comparable, because the effects of the variable are different in the two situations, because measuring units are not coordinate, because other factors modify the effect, or what? Failure to obtain a comparable behavioral effect for a given variable for different tasks raises questions that are very difficult to handle when the interest is in generalizing across tasks

A second danger may be present when equivalent results for a given variable are obtained for two or more different tasks A compelling interpretation of such a result is that the same processes are involved in producing the effect Thus, if distributed practice facilitates acquisition on a pursuit rotor and also on a serial list of verbal units, a ready inference is that the same fundamental process (or processes) are involved This may or may not be true Usually, however, these doubts can be resolved by further research in which further implications of the assumed identity of processes can be tested Indeed, in spite of the dangers and limitations of this method it does not seem unreasonable to suggest that the use of this approach is necessary if unification of the disparate areas of research in human learning is to be achieved

A further way of evaluating commonality of learning in different situations is by what may be called transition experiments An illustration would be the work of Richardson (1958) A descriptive difference between concept formation and rote verbal learning can be stated in terms of the number of identical responses to be associated with similar stimuli Richardson translated this descriptive difference into graded operations so that continuity (or lack of it) between concept formation and rote learning could be studied directly Probably far too little use has been made of transition experiments although there is admittedly much difficulty in working these out for certain areas Still, the ingenious researcher may find extraordinary repayment in devising transition experiments

At quite a different level of discourse is the procedure whereby tasks

and the behavior required of the *Ss* in the tasks, are analyzed along descriptive dimensions with the similarities and the differences being catalogued Thus problem solving and verbal learning are said to differ in terms of the amount of discovery involved in determining an appropriate response, whereas free operant learning and problem solving both involve considerable discovery There are many such descriptive dimensions which might be used (Melton, 1941) At the very minimum, such an approach has clear heuristic value in organizing the tasks used to study human learning But whether or not the descriptive dimensions have psychological or behavioral relevance is a matter for research, and while occasional attempts have been made to determine this relevance within a given area (e g , Riley, 1952), no concerted programs have been undertaken

The Present Approach

The search for continuity by whatever method (and there are undoubtedly others than those mentioned) may, in the writer's opinion, be superficial and perhaps even misleading until the level of analysis in a given area of study reaches a certain stage of sophistication This statement is made from a background of research in rote verbal learning and requires explanation The major purpose of the present paper is to look at the facts and theories of rote verbal learning with an eye toward those that reach beyond the particular situation from which they are derived In a manner of speaking, the purpose of the paper is to view all of human learning through the window of a memory drum To do this in any but a superficial way one cannot look at gross phenomena. To point out that retroactive inhibition occurs in motor learning, in concept learning, and in verbal learning would seem to be of little value We need to know the comparability of the mechanisms involved in producing the retroactive inhibition for the various tasks, we need to isolate the subphenomena to see if the total effects are compounded in the same way How far these analyses must proceed before comparabilities or continuities can be said with confidence to exist is debatable However, one of the points to be made in the subsequent discussion is that the research in verbal learning, one of the oldest areas of research in human learning, has, in the last few years, reached a level of sophistication where we may be in a position to make rather fundamental contact with at least some other areas of human learning

The terms "gross phenomena" and "subphenomena" have been used above Before proceeding it will be wise to give an illustration of how these terms are to be used Consider a transfer situation of the A-B, C-B type (stimuli in two lists different, responses identical) An experimental group learns the two lists, a control group learns only the second The

gross phenomenon in this situation is the amount and direction of transfer which occurs in learning C-B, as determined by the difference in performance of the two groups At the present time we know that this gross effect is constituted of a number of independently demonstrable subphenomena

1 A positive effect produced by transfer of response learning as such Having learned the responses in A B this learning will transfer and the amount of transfer will in turn depend upon certain variables e g., meaningfulness

2. Transfer of response differentiation may occur in certain situations If the responses have interfering similarity relationships, differentiation must be established in order to learn the first list and this differentiation should transfer positively for the experimental group

3 Warm up developed in learning the first list for the experimental group should produce a positive effect in learning the second list.

4 Learning to-learn should produce a positive effect

5 Backward associations should produce a negative effect

6 Others as yet unspecified.

While we can identify these factors as being involved in the gross effect we cannot specify at the present time how they interact or go together to produce the over-all effect Steps in this direction are clearly in order However, even these subphenomena may be further reducible into further subphenomena Learning to-learn must consist of a number of components which, with appropriate research, could be broken down, whereas today it stands largely as an irritating factor to be always considered in the design of experiments So, in fact, when a statement is made (as was made earlier), that the search for relationships among learning demonstrated on different tasks requires a certain level of analytical sophistication in the different areas to expect success, it is difficult to specify just what that level is That we are said to have reached this level for certain phenomena in verbal learning is an expression of confidence undisciplined by any firm criteria

Characterization of Rote Verbal Learning

Rote Learning! Let us imagine some free associations which these two words might elicit from people in psychological and educational circles, restricting the responses to those which meet standards of good taste It is likely that the following would be among the most frequent responses "dull," "Ebbinghaus," "narrow," "verbal learning," "sterile," "nonsense syllable," "memory drum," "serial list," and so on Two notions can be culled from such associations First is the notion that rote learning is closely associated with verbal learning, an association which is quite appropriate The second notion is that rote learning, identified as the classical area of rote verbal learning, is felt to be dull, narrow, sterile, and, in a manner of

speaking, deals with a form of learning that is almost intellectually demeaning. These assumed reactions to rote verbal learning may paint a somewhat exaggerated picture of an attitude toward the area, but most would probably agree that the core of such an attitude does exist. If a number of research areas in human learning were put into the cruel and grinding dimensions of the semantic differential it seems clear that rote verbal learning would come out with a 'bad' profile, perhaps being pressed only by classical conditioning.

Obviously, from the writer's point of view, these attitudes are ill begot. But, perhaps these attitudes are, like so many other attitudes, representative of a cultural lag. Perhaps if the contemporary work in verbal learning were understood by all some change in attitude might occur. Perhaps it is not quite justifiable to view the nonsense syllable as the pedant's playmate. In any event, the position taken here is that the work in verbal learning—rote verbal learning—may stand squarely in the center of all human learning. Research in verbal learning is shooting out phenomena and theories which are touching, sometimes in a very fundamental way, all areas of human learning from simple conditioning to the study of the thought processes. The central nature of verbal learning will not be argued nor pressed, but there is implicit hope that its centrality may emerge obliquely but clearly from the presentation to be made.

The image of a subject in a verbal learning experiment as being a *tabula rasa* upon which the investigator simply chisels associations, and quite against the *S*'s wishes, is archaic. The *S* is far from passive and the tablet has already impressed upon it an immense network of verbal habits. Some of these habits are simple and direct and some are conceptual in their inclusiveness, i.e., they are second-order habits. A more accurate description of the verbal learning experiment is one in which the *S* actively "calls upon" all the repertoire of habits and skills to outwit the investigator. Since the *S* is all too frequently successful in this endeavor, it is quite understandable why some investigators in the field of verbal learning do indeed wish the *S* were a passive and clean slate upon which the associations could be inscribed in a simple fashion. That this is not the case has several important implications, some of which will be mentioned here.

One implication is that in verbal learning experiments we may not be dealing with 'raw' learning, some might say that we never study the formation of associations uninfluenced by associations which the *S* already possesses. Certainly we would approach a verbal vacuum more closely if we spent more of our time studying the verbal learning of young children and less of our time on the college sophomore. But even the young child has verbal associations which will influence the learning of new ones, and in-

vestigations of verbal learning with the preverbal child poses mechanical as well as philosophical problems

Another implication of the fact that we are dealing with organisms with an already established network of habits is that our theories of verbal learning must inevitably explain phenomena which are, in a manner of speaking, built on top of this network

Theories aside, realization of the fact that we deal with nonpassive *Ss* with a vast set of habits may influence the approach we take in experimental studies of verbal learning The preformed habits may themselves be the object of study with further attempts directed toward trying to understand how these preformed habits influence the verbal learning phenomena under study at the moment The fact that in rote verbal learning we must deal in some fashion with preformed habits, and the fact that these habits are of the nature that may influence learning in many other areas lends some credence to the notion that verbal learning is a transition area between simple learning and ' higher ' forms of learning However this may be, let us look at some of the tasks which have been used to provide the empirical base in the area

Tasks

One of the defining characteristics of verbal learning is, obviously, that verbal units must in some way be presented as a part of the learning task When the task does in fact consist of verbal units the identification is made without much ambiguity if the verbal units are symbols of the learner's language But ambiguity may exist in certain situations Thus, if geometrical forms are used as stimulus terms, each paired with a switch that must be closed when the stimulus term is presented, should this task be called a verbal task? The fact is that many *Ss* will apply a verbal label to geometrical forms, and they may even do so for the so-called nonverbal response, i e , the *S* may learn that this stimulus goes with the "sixth ' switch in the series Clearly, some of the studies emanating from the laboratories at Ohio State University (e g , Alluisi & Muller, 1958), and often spoken of as perceptual motor learning studies, may represent transition experiments between pure ' verbal learning and ' pure" perceptual motor learning Therefore, to say that verbal learning deals with acquisition of verbal units can at best be used only as an operational distinction and, if taken too seriously, may actually hinder the development of our understanding of learning processes which might be quite the same for perceptual motor tasks as for verbal tasks

A second specification of a verbal learning task is that multiple associations are required of the *S* This in turn means that multiple verbal units

are commonly used in the task (lists) In very recent years studies have been performed (e g Peterson & Peterson 1959) in which a single verbal unit has been presented and systematic relationships determined for short term memory Thus a single three letter unit is presented to the *S* for a brief period of time and its recall requested at some later point in time The most elementary unit of printed verbal material is the letter The basic spoken or oral unit is less clear, perhaps the phoneme might be considered the unit when oral responses are required In any event if the letter is the basic printed verbal unit then clearly even a single three letter unit constitutes a three unit list With a three unit list at least two associations must be formed or utilized But this analysis may also mask the actual learning process involved since it is quite possible that for a three letter unit the coding process (to be discussed later) makes a single unit of the three letters

With the above two restrictions in mind we may look at some of the tasks which are used in verbal learning studies

Free learning —In pure form this task involves presentation of a series of units to *S*s under instructions to learn the units in any order On successive trials the presentation order is varied from trial to trial A somewhat more restricted procedure would involve a constant order of presentation on each trial but still allow the *S* to recall the units in any order When in addition however a constant order of presentation is used and the same constant order required of the *S* in learning a different task namely serial learning is identified

To those investigators who like to make specific S—R analyses of learning situations the free learning method presents some difficulty The difficulty is in identifying a particular stimulus for the recall of a particular unit At the present time about all that can be said is that the stimulus for the recall of the items is the general experimental context With certain kinds of lists it may be inferred that the recall of one item serves as the stimulus for the recall of another since the items may be known to be strongly associated But of course the ambiguity still remains as to the stimulus for the recall of the first item It should be clear that inability to identify a specific stimulus for each item does not prevent the determination of lawful relationships for free learning Indeed as will be seen later the free learning method has great value in studying certain subphenomena involved in the over all learning process

Serial learning —In this task the units are presented to the *S* in a constant order on each study trial and he is required to learn them in the order presented In its pure form the units are discrete and the fixed order of presentation is essentially determined on a random basis This method

dating back to Ebbinghaus, has been used in countless studies Serial learning produces the bowed serial position curve, a phenomenon which has tantalized theoreticians for many years Again, investigators who like to make precise S—R analyses have, in recent years been shying away from this method since ambiguity exists in identifying a specific stimulus for a specific response It has probably been implicitly assumed by many that the stimulus for a response in serial learning is the immediately preceding term However, recent considerations (Underwood 1963) suggest that this assumption cannot be maintained with any degree of confidence The facts seem to indicate that we cannot specify just what the stimulus is for a given response in serial learning

Progress in a science may be identified in terms of the capacity of the investigator to break down gross phenomena into subphenomena Regardless of the theoretical orientation involved, in the study of learning this implies in part stimulus control and stimulus analysis It therefore seems likely that the frequency of use of serial learning as a standard vehicle for studying verbal learning will further diminish unless a more precise analysis of stimulus and response functions in this task is accomplished soon It is interesting to speculate on the cause for the drop out of tasks in other areas of human learning, tasks which once were heavily used For example, the great flood of studies using the pursuit rotor which occurred in the postwar years has essentially dried up One possible reason for this is that the pursuit rotor is a very difficult task to break down into subcomponents for which stimulus response functions can be specified Investigators tend to drop a task when it will not give them the analytical power characteristic of the field as a whole

Serial learning with syntactical ordering of units—In this task words vary in form class (nouns verbs, and so on) and the ordering of the words is such as to conform to the rules of grammar The *S* is required to learn the words in the order presented The experimental unit for such a task may be a sentence and the acquisition of several such units (perhaps forming a paragraph) may constitute the task Of course, the scoring of learning may occur at several levels, such as number of words number of idea units and so on Although the items presented in this task must be words in order to realize the full effect of syntax, similar effects can be produced with nonsense words Thus, Epstein (1961), by changing endings on various nonsense words to suggest units of various form classes in appropriate ordering was able to show that learning was positively influenced

It is perhaps needless to say that this learning task appears to involve properties which none of those traditionally associated with rote verbal learning have Furthermore, the task has not been used in many studies of

learning Yet it seems likely that in the future years this task will provide an important point of contact between verbal learning, psycholinguistics, and other disciplines interested in the study of language behavior

Verbal discrimination —In verbal discrimination learning the *S* is presented a series of *pairs* of units The investigator arbitrarily designates one member of each pair as the correct member, this designation being random with regard to position or any other characteristic of a conceptual nature which might obtain across pairs Initially the *S* guesses one or the other units of a pair and is told in some manner whether his guess was right or wrong The order of the pairs in "pure" verbal discrimination learning varies from trial to trial If the order remains constant a multiple unit two choice maze is simulated

Verbal discrimination learning emphasizes recognition learning rather than recall learning In verbal discrimination learning the subject does not have to learn the responses per se as in other forms of verbal learning, but clearly a discrimination of some kind must be developed between the units However, as a task it provides a fairly close approximation to the learning required in distinguishing among stimuli of a paired associate list and therefore should prove of considerable value in future attempts to break down stimulus functions The fact is, however, that verbal discrimination learning is not frequently used as a method for studying verbal learning

Paired associates —Pairs of verbal units are presented to the *S* just as in a verbal-discrimination task Here, however, the left hand member of the pair is designated the stimulus term, the right hand member the response term The *S*'s task is to learn to be able to recall the response term when the stimulus term is presented alone The pairs are commonly presented in different orders on each trial This task is by far the favorite of contemporary investigators since the stimulus for a given response can be fairly precisely specified However, even this specification is not without ambiguity as later discussion will point out

The above five tasks may be taken as modal tasks for the study of verbal learning None of the procedures is fixed or immutable, nor should they be Thus, there may be more than two choices in a verbal discrimination task, the paired associate list might be presented in constant order, the pairs in a paired associate list might have minimal syntax (e g , the stimulus terms might be nouns the response terms verbs) and so on

All of the exposition thus far had as its purpose the setting of a background for the remainder of the paper We may now turn to an evaluation of certain phenomena, derived from the study of rote verbal learning, which are potential phenomena for bringing together some of the results of the diverse investigations of human learning

Factors in Response Learning

The reference point for most of the remaining discussion will be paired-associate learning. As the first step two gross stages in the over-all learning may be identified. There is first response learning, the *S* must be able to recall the responses as such. The second stage is the associative stage wherein each response term must be associated with a specific stimulus term (for a more extended discussion of these two stages see Underwood & Schulz, 1960, pp. 92–94). This section is concerned with factors involved in response learning. It should be clear that the intent is not to summarize the effects of all manipulable variables, rather, the intent is to treat those phenomena associated with variables which appear to have some potential for being related to phenomena occurring in other areas of research on human learning.

Acquiring a Difficult Trigram

Assume that VXK is one of several response terms in a paired associate list. If the stimulus terms have low interstimulus similarity, it can be shown that much of the over-all learning time involved is that of being able to recall the response terms correctly. For an equal number of common three-letter words response learning is of minimal importance. Let us look at some of the factors which may be involved in acquiring a difficult trigram.

Specific interference.—First it may be noted that the matter of merely acquiring the letters as such is of no importance for the young adult learner. The usual *S*, having had 15 years of experience with reading and writing can identify and produce the letters of the alphabet. So, it is the appropriate sequencing of the letters which makes the acquisition of the trigram a difficult learning task. In the letter association tables (Underwood & Schulz, 1960) the letter X is rarely given as a response to the letter V, and K is rarely given to VX. Furthermore, these letters rarely if ever occur in this order in words. Thus, the first factor to consider is that there is little if any positive associative connection between the letters, there is little positive effect from previous learning. In this sense, the learning of this trigram by a college student might be likened to a kindergarten student learning the word CAT. However, the *lack* of strong associations between letters for the college student does not seem to be able to account for the difficulty experienced in learning a difficult trigram. The college student, unlike the child, does have strong associations to each letter, that is, other letters are strongly associated with V, with X, and with K. This is not to imply that young children have *no* associations between letters but only that the associations are likely to be weaker than those of the college student.

Indeed, Mrs Sheila Keppel (unpublished) was able to obtain quite reliable letter associations from second grade students Furthermore, the frequencies of particular letters given to single letter stimuli correlated between 34 and 63 with the frequencies of the letters elicited from college students For two letter stimuli, the correlations ranged from 36 to 85 So, it is clear that certain sequential letter habits are built up early in the development of a written language, but for children it is likely that the strength of these associations is less than for college students The implication of this is that the strong alternative associations in the repertoire of college students produce interference in the acquisition of this unusual letter sequence Several points about this interference need elaboration

1 Overt interference, such as giving a wrong second letter, is almost completely limited to letters in the list Letters which are not in the list appear so rarely that reading errors may be responsible for them This selector mechanism, as it has been called (Underwood & Schulz, 1960), is extremely powerful and seems to be characteristic for all verbal learning tasks In a recent unpublished study *Ss* learned 16 pairs of common words for 15 trials The 30 *Ss* involved made 1,424 overt errors, only one of these errors was an importation, the response being "yellow" when the correct response was "canary" Thus, in spite of the fact that all of the words were common, and all would have strong associations, only one of these associates appeared in the overt errors produced by the *Ss*

The failure of strong associates which are not in the list to intrude does not mean, of course, that they do not contribute to the interference However, it is likely that if a verbal unit not in the list is strongly associated with one in the list, the amount of interference it produces is less than for a unit which is equally strongly associated and is in the list

2 If the bulk of the interference is produced by associations with other letters in the list, it is clearly a negative transfer paradigm in which there is a repairing of letters so that pre-experimental associations are inappropriate This situation is very comparable to perceptual motor studies (e g, Fitts & Seeger, 1953) where variations in the compatibility between the experimental requirements and population stereotypes (pre-experimental associations) affect the rate of learning

3 How does the *S* ever learn a task where such strong interfering associations are present? There is reason to believe that a process of extinction occurs, and this is one point at which verbal learning seems to mesh with the area of classical conditioning There is independent evidence for an extinction like process (Barnes & Underwood, 1959), the evidence is also available for a spontaneous recovery process (e g, Briggs, 1954) and the use of both of these concepts is occurring in forgetting theory (Postman, 1961) Whether or not extinction of conditioned responses and extinction

of verbal associations are based on the same processes is not clear For example, not all associations in a list may be extinguished and the reasons for this are obscure A sufficient number of extinction trials in classical conditioning usually will bring the response potential to zero In any event, this area of contact between verbal learning and classical conditioning seems sound enough to merit transition experiments designed to get directly at the subphenomena involved in both

Coding —In developing an association between two units on a memory drum we are sometimes inclined to speak as if the *S* learns an association between the two units as they exist 'out there' on the memory drum Of course this is absurd Both terms must have some representation in the mind of *S*, and the term "mind" is used here to indicate whatever neural-cortical chemical mechanisms are involved These representations are, of course, responses in the strict sense of the term What is the representation for a trigram such as VXK? The instrumental response we require of the *S* is that he spell the unit We may think, therefore, that the instrumental response is in some manner isomorphic to the representation of the trigram in the mind of the *S* However, there is every reason to believe that in many instances the *S* codes a response term (just as he may a stimulus term a matter which will be discussed later) What the coding is presumed to do is to provide a more readily remembered unit than would be the case if the memory unit consisted of a representation that is isomorphic to the unit as presented on the memory drum Thus, there is a discrepancy between the unit of memory and the objective response term When giving the instrumental response, therefore, there is a decoding process

Is a coding process of the response term reasonable? Everything we know would say "yes" In the first place *Ss* are extraordinarily acute in making judgments about relative rates at which they will learn various units Long experience with verbal units apparently provide them with the critical criteria for making these judgments Indeed, the correlation between 'learnability' ratings and actual learning rate is so high that if the interest of an investigator is only in the difficulty of individual items it would essentially be a superfluous step to obtain learning measures One of the criteria which *Ss* use in rating the learnability of an item is how easily it is pronounced Pronunciation is a coding process for it reduces the number of discrete memory units involved Of course the *S* in making the instrumental response must decode, but if the sound unit or units serving as the memory unit is standard, hence well practiced, this decoding adds little to the learning process How much decoding occurs in a difficult trigram such as VXK is not known The very fact that it is difficult may indicate that little coding of a facilitatory nature may occur This problem is clearly researchable For example, if relatively difficult trigrams were presented, which, when

the letters were rearranged spelled a familiar word, would the coding process "change" the trigram into a word For example, suppose the trigram presented a subject was GDO Is this trigram coded as GOD or DOG with additional learning to implement making the correct instrumental response? This whole matter of the relationship between the response term per se and the representation of it in the actual learning is mentioned here as background for subsequent discussion on both stimulus-term and response-term coding

Motor involvement —In most verbal learning tasks little attention is paid to possible motor components in the instrumental response Thus, on a rating scale of amount of motor involvement, verbal learning would be rated much lower than, say, pursuit rotor learning Whether or not the supposed small amount of motor involvement in producing a difficult trigram is of psychological importance is not known It is known that the mere speed of spelling difficult letter sequences is slower than the spelling of less difficult sequences (Waters, 1939) The slow learning of a difficult trigram could in part be due to motor conflict since the sequence of sounds are new to the *S* The difficulty young children may have with a certain sound or sequence of sounds indicates that there may be some fairly precise motor skill learning involved Certain foreign language sounds provide very difficult motor patternings for the adult attempting to speak the language for the first time It is possible, therefore, that appropriate studies would show reminiscence in verbal learning comparable to those shown with the pursuit rotor or other tasks commonly identified as emphasizing precision of movements At the present time, however, there seems to be no data in verbal learning which provide a compelling bridge to motor learning

Second-order habits —The discussion proceeds with factors judged to be important in learning a difficult trigram The intent is to suggest the possibility that the trigram, such as VXK, is not only difficult for the various reasons already noted but also because certain second-order habits may be evoked in the learning situation We have noted earlier that letter-association data predict that VXK would be a difficult trigram to learn and that in fact it is Thus far the interference factors said to be involved in learning such a trigram have been specific in that the interference comes from associations with other specific letters which have been previously associated with each given letter The letter-association data suggest a further potential source of interference If letters are classed as vowels and consonants, the letter-association data indicate that the likelihood of giving a member of one or the other classes to a specific letter stimulus corresponds very closely to the sequential frequencies of consonants and vowels in words Thus when a single letter stimulus is a consonant, 62% of the [illegible] vowel and 38% with a consonant These values

correspond exactly to the consonant-consonant and consonant vowel frequencies in words. The correspondence in the case of a vowel stimulus is not quite so high but is still rather striking (see Table 39 of Underwood & Schulz, 1960). When the stimuli are two letter units with the *S*s asked to give a third letter, the correspondence between sequences in words and the response given by the *S* maintains itself.

These facts suggest that responses in the letter association procedure may be due in part to second-order habits relating sequential frequencies of consonants and vowels without regard to the specific letters involved. Or to put this another way, environmental regularities of classifiable specific events have induced second-order habits or concepts. For reasons explained by Underwood and Schulz (1960), the letter association data as we viewed it did not require the assumption of these second-order habits: most of the facts could be accounted for by assuming only that specific associations were involved. Yet, there were some findings which could not be accounted for by assuming that differences in environmental frequencies of sequences were reflected only in the specific associations. For example, H occurred very infrequently as a consonant response in spite of its high frequency of occurrence following specific letters in words. Specifically, H follows T with very high frequency in words, yet only 1 *S* out of 273 gave H when T was the stimulus. Thus, in spite of the fact that the data do not firmly justify it, an assumption of second-order habits in letter sequences will be made. Two implications of this assumption will be discussed.

1. Second-order habit systems may produce interference in learning certain letter sequences, perhaps even three letter trigrams consisting of three consonants. That is, since a sequence of three consonants occurs infrequently in the language, other second-order habits (consonant consonant vowel) may introduce interference by introducing a tendency to give a vowel as the third letter. Such an interpretation may be a little forced in that there are at least two unknowns. First, we do not know in the learning situation how much of a sequence is necessary to activate the second-order habit. In the letter association data, responses indicating an alphabetical habit were of much higher frequency with two letter stimuli than with a single letter as stimulus. For example, C was given as a response to B by only 8 of 273 *S*s, but C was given as a response to AB by 135 *S*s. Thus this habit (which is not here considered a second-order habit) is more likely to be induced by a two letter sequence than by a single letter. Such a result is also consistent with the finding of Howes and Osgood (1954) on the effect of context words on word associations (see particularly their Exp. 2). If second order consonant vowel sequence habits are elicited in learning a trigram we cannot be sure of the potency of such habits; it may take more than two letters to bring them to full activation.

The second consideration is whether or not the selector mechanism, which is assumed to reduce interference from specific associates which are not represented by the response unit in the list being learned, may also operate to "curb" second-order habits When all the response terms are consonant syllables, perhaps no vowel habits are elicited In word-association data the responses to a single word rarely follow grammatical order (a second-order habit) Thus, to the stimulus "dog" a verb such as "ran" or 'barked' would rarely be given If the stimulus consisted of "The dog," the word association would undoubtedly change to "ran" or 'barked" and the activation of the second-order habit might well influence responses to subsequent single word stimuli In a similar fashion different vowel and consonant sequences among the response terms may be required in order for the second-order habits to be activated Again, this problem can be answered empirically by constructing lists of response terms in which some items correspond to the sequential patterns of vowels and consonants in words and in which some do not These latter items should be more difficult to learn than when all items in the list correspond to the same sequential pattern

2 The above discussion assumes that second-order habits reflect with considerable accuracy the frequency of different sequential events which occur in the environment. The development of such habits must be reflecting very much the same habit mechanisms being studied in the area called probability learning In the probability learning studies the investigator presents events of differing frequencies to the *Ss* and studies the degree of correspondence manifested between these frequencies and the choice behavior of the *S* Any given choice cannot be adequately accounted for by the immediately preceding stimulus, the choice is determined in part at least by a *series* of preceding stimuli. This is oversimplified of course, but it does seem that the two areas make contact at this point It is also possible that the second-order habits, presumed here to be operating in the learning of letter sequences may have parallels in the memorization of sentences where syntax habits are important, these syntax habits also producing interference when the sequence of words form a jumbled sentence

Similarity and Response Learning

Free learning, in spite of its lack of amenability to S—R analysis, provides an ideal method for studying factors involved in learning responses as such In this section we will deal with the fact that similarity among a group of units will facilitate the free learning of them The major point to be made is that in the manipulation of intralist similarity in verbal learning we are essentially dealing with concept formation Thus, the section will be an amplification of Miller s (1958, p 490) statement

> Although some psychologists seem to hold the opinion that concept formation is merely a complicated kind of verbal learning the present analysis suggests the converse—namely, that verbal learning is merely a simple form of concept formation In either view however it seems likely that the same problems will arise in both areas and that one will not progress without the other

Miller's work leading to the above statement used sequences of from 4 to 7 letters, with 9 sequences constituting the list Only four consonants were used in making the 9 units In certain lists the consonants were ordered randomly, in others they were ordered so that certain patterns of letters were repeated in the various units If these patterns were observed by the *S*, as they apparently were, the learning of such units should occur more rapidly than the learning of the random sequence units The outcome of the experiment clearly supported the notion

In the Miller experiment the basic patterns, which may be considered concepts, were both learned and utilized in the experimental situation The utilization of concepts which have been pre-experimentally learned is believed to be quite convincingly demonstrated in the following unpublished experiment

Conceptual similarity and free learning—Four lists of words, shown in Table 1, were used in the study Lists 1 and 4 have low interitem similarity

TABLE 1

LISTS USED TO STUDY FREE LEARNING AS A FUNCTION OF CONCEPTUAL SIMILARITY

List 1	List 2	List 3	List 4
apple	Bob	France	daisy
football	Bill	England	wall
emerald	Joe	Russia	bee
trout	John	Germany	second
copper	rabbi	bluejay	knife
theft	priest	canary	bus
hat	bishop	sparrow	geology
table	minister	robin	maple
cruiser	cow	measles	arm
trumpet	horse	mumps	hammer
doctor	dog	polio	salt
head	cat	cancer	tent
wine	rumba	nitrogen	cobra
blue	foxtrot	oxygen	mountain
gasoline	tango	hydrogen	window
cotton	waltz	sulphur	rain

Lists 2 and 3 consist of four sets of four words all of which were high frequency responses in the tables for responses to category names as given by Cohen, Bousfield, and Whitmarsh (1957) These four lists were presented to 37 *S*s In presenting Lists 2 and 3 the ordering was not as shown in Table 1, rather, in each block of four items each of the four categories was represented once, but the ordering of the four items within a block was random

The rate of presentation of the items was 5 sec, during which period the experimenter spoke the word twice The instructions gave no hint that concepts or categories were involved in Lists 2 and 3, and since the lists were presented in the order 1, 2 3, and 4, the *S*, after having had List 1, would have no expectation that List 2 would consist of four sets of four words falling into the same category Each list was presented once and the *S*s were allowed 2 5 min to write down all the words possible Instructions made it clear that the words could be written in any order

The mean numbers of items recalled were 11 08, 14 57, 14 86, and 11 35, for Lists 1 through 4, respectively Summing the recall for the two high similarity lists for each *S* and the two low similarity lists for each *S* shows that 36 of the 37 *S*s produced more correct responses for the high similarity lists than for the low Of the 74 recall protocols for high similarity lists (two for each *S*), 28 (38%) showed perfect recall, for the low-similarity lists only 2 (3%) perfect recalls were observed Thus, these gross results basically confirm a number of other studies using varying materials (Deese, 1959, Horowitz, 1961, Underwood, Runquist, & Schulz, 1959) Therefore the problem of interpretation is the matter of concern The position taken here is that the critical unit of memory involved in the present study is the concept or category name Additional evidence is necessary to make the point

1 In recall of the high similarity lists clustering was nearly perfect Only five of the 37 *S*s might be said not to have shown extreme clustering The other 32 *S*s in general produced recall protocols in which all four items in a category were recalled together then another four, and so on The *S*s were not told the number of items in the list yet it was clear in many of the protocols that the *S* knew there were 16 items and four instances of each of four concepts No *S* ever gave five words from a concept In the 74 recalls only three showed a failure to recall any word from a category That is, these three protocols showed perfect or nearly perfect recall for the 12 units forming three categories but no recall for the fourth In several of the protocols in which 15 items were given correctly, a blank space was left for the fourth item in the category No *S* ever wrote down less than three items from a given category

2 The number of intrusions was identical for the high and low similarity

lists A total of 16 was observed in each set of 74 protocols Again, the precision of the selector mechanism is apparent Of the 16 intrusions for the high similarity lists, 14 were "good" intrusions in the sense that they were members of one of the four categories Only the word SAMBA occurred twice, replacing RUMBA in both instances One of the intrusions was TABLE which was a member of List 1, and the sixteenth intrusion was not readable

These data strongly suggest that the basic memory unit involved was the category name Let us take the category of the four dances When the first of these is presented, say, FOXTROT, it seems inevitable that it must elicit an implicit response, DANCE When WALTZ is presented it too should elicit the same implicit response, and so on When several words elicit the same response it becomes a more frequent response than any of the words considered separately Memory for categories was almost perfect in that of the 296 different opportunities to remember categories, there was a failure in only three cases

What about an alternative interpretation which simply says that if the S remembered four specific items, one from each category, and then free associated to these the present results would be obtained? Word associations to each of the items used is not available, but it seems quite unlikely that if they were available they would predict the present results That some relatively free association within the category might have occurred cannot be denied, but even this was probably minimal since relatively few intrusions occurred and since a number of protocols showed that the S knew that one more word was required for a given cluster but none was given This failure in the latter instance could not possibly be because the S could not think of another unit that fitted the category, rather it indicates a clear editing process Thus, while the category name or concept played a fundamental role in the recall process, it is also evident that the S carried considerable knowledge about the particular entries under each concept

Bousfield (e g , 1953) has argued for the importance of second order habits in memory, the work of Mathews (1954) also supports the importance of conceptual memory The present study would give such arguments full support In short, it seems inevitable that when verbal units evoke converging implicit responses a facilitative feature for the memory of particular units results Verbal learning does indeed overlap the study of concept formation and utilization

Factors Influencing Associative Learning

In the previous section factors were discussed which appear to be important in learning responses as such, with the reference point being a paired

associate list In this section attention will be directed toward factors involved in attaching a particular response term to a particular stimulus term The interest remains primarily in attempting to identify phenomena which have generality beyond the area of rote verbal learning

Stimulus Selection

In the previous section comments were made about response coding The first consideration is to recognize that just as in the case of a response term, the stimulus term that appears on the memory drum does not become associated with the response term The association is initiated by whatever representation the stimulus term produces in the mind of the *S* This stimulus term representation may, theoretically at least, be associated directly with the response term representation although as later discussion will make clear there is considerable evidence that it may not be quite that simple, at least with certain materials

In the usual paired associate procedure the *S* must produce the response as represented to him While coding and decoding may be involved in this, he nevertheless is forced to produce the response In the case of the stimulus term, the usual paired associate procedure does not require its production Thus, the *S* may code the stimulus term without ever having to decode it. The facts seem to be that in certain situations the *S* selects aspects of the stimulus situation and this selected portion becomes the basis on which the association between the stimulus and response term is established Evidence for such selection has been reviewed elsewhere (Underwood, 1963) and will not be repeated here It may be mentioned that to recognize that stimulus selection may occur places quite a different perspective on certain experiments in rote learning, such as those dealing with context effects on retention

The fact that stimulus selection may occur in another illustration of how we cannot conceive of the adult *S*s in our experiments as being passive organisms upon whose minds we inscribe associations They bring to the situation well-developed habits of learning and they impose these habits on the situation As noted earlier, *S*s can make amazingly acute judgments concerning the ease or difficulty of learning verbal units Thus, if the *S* is faced with a paired associate task in which each stimulus term consists of two elements, one a common three letter word, and the other a difficult trigram, it can be shown that most of the associations develop between the word stimuli and the response term and not between the trigram and the response term (Underwood Ham & Ekstrand, 1962) To say that the *S* knows he can learn faster by using the word rather than the trigram may not be completely satisfactory It may be more appropriate to say that the word elicits an immediate and stable representational response whereas the

trigram does not In any event, it is quite clear that what has been here called stimulus selection occurs

The problem of stimulus selection in simple discrimination learning in animals has been a persistent one although often ignored for it is likely to produce pesky concepts like attention Certain learning theories have been criticized recently by Dodwell (1961) for failure to consider the cue value or coding aspects of various stimuli That there are at least preferences for dealing with certain stimuli over others has long been recognized in the area of concept formation and problem solving Thus, when different sets of stimuli have logically equal probabilities of producing solution to a problem it is found that there are biases toward dealing with certain sets of stimuli in preference to others A recent study of concept formation by Shepard, Hovland, and Jenkins (1961) clearly demonstrates stimulus selection That such selection can be developed experimentally for different dimensions has been shown by Eckstrand and Wickens (1954) So, the problem of stimulus selection is a fairly universal one

Preference vs ease of learning—The above discussion indicates that *Ss* may prefer to deal with certain stimuli than with others One of our stated implications of stimulus selection in verbal learning was that the *S* selects a portion of the stimulus complex as the functional stimulus because past experience has shown him that learning will occur faster with one particular component than with another This may be true in some situations but it need not be, indeed should not be, taken as a general principle Some scattered evidence suggests that we must be very careful in not confusing preferences with ease of learning, a distinction which is another side of the performance learning distinction Three sets of evidence will be mentioned

In a recent study (Underwood, Ham & Ekstrand 1962) compound stimuli were used in a paired associate list One of the components of each compound was a common three letter word The other component was a rectangular frame surrounding the word, each rectangle being of a distinctive color The *Ss* learned the list and then some were transferred to a list in which only the colored frames were stimuli Performance on the transfer task was significantly better with words as stimuli than with the color frames as stimuli This led to the conclusion that more *Ss* selected the words as the functional stimuli than selected color One interpretation might be that words were selected more frequently because the experience of the *S* led him to believe his learning would be faster with words than with colors If this is so, it did not work out in practice for when new groups of *Ss* were used in which either the color frames or the words were used as stimuli, the learning proceeded at about the same pace

A second illustration may be obtained from the work of Postman and Riley (1957) In the lists used by these investigators the units were three

digit numbers or nonsense syllables If the stimulus terms and response terms were both numbers, or both syllables, they were called *like* pairs If the stimulus term was a number and the response term a syllable, or the reverse, they were called *unlike* pairs Some lists were constructed in which all pairs were like pairs and other lists were made up in which all pairs were unlike pairs In addition, mixed lists were used in which half the pairs were like and half unlike The results of this experiment show that in a mixed list consisting of four like items and four unlike items, the like items produced a criterion of perfect learning (all correct on a single trial) significantly sooner than was the case for the unlike pairs However, when the entire list was made up of like pairs or of unlike pairs, there was no difference observed in attaining one perfect trial The same results were obtained in a subsequent experiment (Postman, personal communication).

These results suggest that the *S*s had a preference for dealing with like pairs but that in fact such pairs are no easier to learn than unlike pairs When the *S* is given a mixed list consisting of like and unlike pairs he cannot learn all items on a single trial, in a sense, he has to start somewhere and apparently the bias is toward starting with like pairs

In still another study (Karwoski, 1931), 12 two-digit numbers were presented to *S*s The variable was the size of the print, being small, medium, or large print sizes When the list consisted of items having all the same print size there was no difference in rate of learning When, however, a mixed list was used in which four numbers appeared in small print, four in medium, and four in large, the large numbers show a clear initial advantage in performance over the small and medium numbers

Two conclusions from the above facts may be drawn First, in all areas of learning where applicable we must draw a distinction between preferences for dealing with particular stimuli and the rate at which those stimuli may enter into associations The second conclusion deals most directly with verbal learning It is a fact that in at least one study of transfer (Twedt & Underwood, 1959) the same results were obtained with mixed and unmixed list designs The results we have just cited suggest that the finding with transfer designs may not have generality When tasks variables are manipulated the mixed design may lead to erroneous conclusions about rates of learning when what in fact is being measured is preference for dealing initially with certain stimuli over others

Mediation

Mediation holds a firm grip on many areas of learning, and verbal learning is no exception To a large extent, mediating mechanisms are theoretical constructs when applied to animal learning (e g , the fractional antici-

patory goal response), and they retain much of the same status when applied to human learning of various kinds

In paired associate learning there is little doubt that simple mediation occurs Subjects too consistently report the use of verbal mediators to link the stimulus representation and the response representation to deny the basic processes Thus, if a pair in a list is DOG—9 the *S* may use a mediating term CAT, the association running from DOG to CAT to NINE since some already pre established associations prevail between DOG and CAT and between CAT and NINE Or consider a transfer situation in which in the first list the S learns DAX—ICY and then in the second list is asked to learn DAX—FRIGID In one experiment (Barnes & Underwood 1959), 94 out of 96 *S*s reported the second list items were learned by using the first list response as a mediator so that the chain in the second list performance was DAX—ICY—FRIGID All other measures supported the reports given by the *S*s (e g , performance on the second list was essentially perfect on the first anticipation trial) In short as far as simple mediation is concerned there seems to be no doubt that there are reportable and measurable counterparts of the theoretical notions of mediation Thus in verbal learning it is clear that at least under favorable conditions the association between the stimulus and response terms is formed by activation of already established associations That all expectations concerning the effects of mediation do not work out with a powerful effect has been clear in some of the work at the University of Minnesota (e g Horton & Kjeldergaard 1961) where a great deal of the research on mediation in verbal learning has been centered

The heavy emphasis in recent years on mediation has seemed to leave little room for new or raw associative learning With the young adult do we ever study the development from scratch of a direct association between the representation of the stimulus term and the representation of the response term? If we are to take *S*s reports seriously the answer is yes Subjects will report that the stimulus and response term being paired together on several trials finally just go together They report no mediating term of any kind The present writer is inclined to accept these reports as representing a true state of affairs It almost seems necessary that some such process be used to account for the development of associations between the letters of difficult trigram response terms The present writer is quite willing to accept the idea that some associations in verbal learning experiments are developed by mere contiguity of stimulus and response terms While accepting the fact that our *S*s have enormous networks of associations it does not seem likely that all of the learning we observe is simply the utilization and strengthening of old associations there must be

some new learning In any event, there does not seem to be evidence to justify letting the notion of mediation get out of hand

Verbal mediators are being used in explanatory accounts of learning in other areas Goss (1961) has examined these and so they need not be reviewed here It is sufficient to say by way of summary that not only do mediators play a role in associating stimulus and response terms in paired associate lists but also that the mechanism per se apparently has wide applicability to other areas of research in human learning

Intralist Stimulus and Response Similarity

In considering how an association is established between the stimulus and response term in paired associate lists probably no variable looms so important as intralist similarity We have already seen how intralist response similarity may facilitate response acquisition, but the very mechanism by which this facilitation takes place will retard the development of specific stimulus response associations Likewise, similarity among stimuli will retard the development of specific associations

Basically, intralist stimulus similarity must produce an inhibitory effect because the initial representational responses to stimulus terms are similar if not identical Thus, the *S*'s difficulty comes about in developing a unique representational response to each term or a unique mediational response If the similarity is great enough the necessary discriminations are some times impossible For example, we have used paired associate lists consisting of only four pairs of consonant syllables but having high interstimulus, high interresponse, and high stimulus-response similarity (Underwood & Richardson, 1957) An appreciable number of college students were unable to learn such a list in 150 trials

The manipulation of intralist similarity in verbal learning results in changes in a number of subphenomena Some of these exert a positive effect on learning i e , as similarity increases a facilitating effect on response learning is produced The weight of the subphenomena is, of course, generally negative but in view of some of the occasional puzzling results (Underwood, 1953), it should not be assumed that the negative effects will always be greater than the positive The cataloguing of these various subphenomena need not be given here since most of them seem to be limited to the manipulation of similarity in verbal learning studies An exception will be mentioned later

In one form or another similarity has been manipulated in tasks across all areas of human learning A coordinating concept which appears consistently in the literature is that of *generalization* (Mednick & Freedman 1960) Yet the unifying nature of this concept may be illusory since it has many meanings each representing some distinctiveness in the assumed mod-

of operation In addition, if generalization is assumed to be the mechanism by which the negative effects of intralist similarity are produced, an important issue concerns how this generalization is reduced so that an *S* can in fact learn the task presented him Both of these matters require further discussion

Conceptualizations of generalization —In the following illustrations the references will be to various ways of manipulating stimulus term similarity and the corresponding ways in which generalization may operate Without further reference to the matter, it will be assumed that the same forms of operation may occur when response term similarity is manipulated

The classical conception of generalization is sensory in nature Thus, visual stimuli of nearly the same wavelength may elicit the same response, or forms of nearly the same shape may be responded to in identical fashion With strictly verbal materials the nearest approximation to this would occur with low meaningful consonant syllables Thus, supposing two stimulus terms were WJG and WGJ Not only do these two terms have perceptual similarity but a spelled representational response would produce high overlap for the two units Both the visual and auditory responses to the two units would converge to produce a highly similar representational response Given this situation, it is quite possible to apply the notion that as a response becomes associated with one or the other, there is a spread of excitatory or associative strength to the other If different responses are required to the two stimuli, interference in learning should result The representation of the similarity along a unitary dimension is probably not possible but that a complex dimension of sensory similarity exists in this situation seems quite reasonable

Next, consider two words which have meaningful similarity, say, ICY and COLD It seems clear that these two units cannot be considered very similar along sensory dimensions By what mechanism would such units produce interference in learning if different responses had to be associated to them?

The mechanism could be conceived of in exactly the same fashion as for sensory generalization Each unit may elicit the same response which in this case may be of an affective nature, or, insofar as these items might elicit common responses on various scales of the semantic differential we would clearly have a highly similar and perhaps complex mediator produced by the representation of these stimuli This seems to be close to what Osgood, Suci and Tannenbaum (1957) mean by meaning Thus while the nature of the dimensions of similarity differ from those commonly associated with sensory generalization and is one step removed from the representational response, it nevertheless could have the same mode of operation

There is another method by which intralist similarity may produce inter

ference for two such words as ICY and COLD Highly synonymous words have high associative connection (Haagen, 1949) Thus, if an association develops between ICY and its response term, when COLD is presented the sequence of associations may run from COLD to ICY to the response term for ICY In the usual sense, this would not be called sensory generalization

Still another mode of operation could be assumed which is somewhat more complex but combines both mechanisms already mentioned Why do two words, similar in meaning, have high associative connection? Perhaps the association is not direct Let it be assumed again that ICY develops some associative strength with its response term in a paired-associate list Now, when COLD is presented, the association may run from COLD to a mediating response which leads to ICY and then to the response term associated with ICY

Consider another case of similarity where assumptions of a common mediator seem a little more firm than is the case where synonymous words are involved In an analysis made earlier of the recall of lists (Table 1) having sets of items with conceptual similarity, it was argued that the category name mediated the recall In a further experiment (Underwood & Schulz, 1961) these kinds of material were used in paired associate lists and learning was studied over several trials In one list four items representing one concept were stimulus terms and four items representing a different concept were paired as response terms Thus, four animal names were paired with the names of four countries There were four such sets of pairings in the list for a total of 16 pairs The learning of such a list proved extraordinarily difficult As judged by the errors, the learning of the concept pairings was almost immediate, 97% of the overt errors, of which there were 1424, were "good" errors, e g , a country name was given to an animal name But, discriminating which particular animal name went with which particular country was very difficult It was as if the gross mediating responses (category names) were so strong that they interfered with the development of more precise discriminations While further study is needed to fully understand this situation it does seem apparent that a common mediator—a concept—is involved in a negative manner This might suggest in turn that the notion of a common mediator for meaningful similarity (the ICY—COLD illustration given earlier) would be quite tenable, but it may be a mediator without a verbal counterpart such as seems to be the case for conceptual similarity

The above discussion of various ways by which generalization may be conceived results in some common features Looking across all the ways by which similarity may be manipulated, it may be suggested that similarity must produce its effect because of sensory similarity of the representational response or because of the similarity of the response (the mediator) to the

representational response But even if these basic commonalities are accepted, there are still different ways of viewing the detailed mechanisms (e g, Bousfield Whitmarsh & Danick 1958 Dicken 1961)

Generalization reduction—Given an interference produced by similarity and interpreted as being due to generalization how does the subject rid himself of these interferences We have very little information on this When the interference is produced by the similarity of representational responses it seems reasonable to expect that stimulus selection occurs to reduce the similarity Another possibility would be in terms of reduction of the range of the gradient of generalization by nonreinforcement of errors For a number of reasons including the lack of evidence for spontaneous recovery of generalization (Underwood, 1961), it seems unlikely that this interpretation will be found appropriate for verbal learning When similarity produces its effect through similarity (or identity) of mediators the mediating associations may go through an extinction process Generally speaking however, we know very little in a substantial way about these processes if they do indeed occur

As noted earlier similarity in some form is a variable that is manipulated in all areas of human learning The resolution of the various problems noted above will probably serve as a very useful organizing technique for bringing more unity into the learning observed in the various situations in which human learning is studied

Summary and General Comments

In this paper certain phenomena derived from research on rote verbal learning were analyzed in terms of their relevance to other areas of research on human learning when these phenomena appeared to have some counterpart in the other areas In an earlier draft of this paper a section was devoted to backward associations It is a fact that in paired associate learning the associations which develop are to a large extent bidirectional Not only will the stimulus term elicit the response term but the response term will elicit the stimulus term While it is difficult to conceive of bidirectional associations for certain tasks in some areas of research (e g eyelid conditioning pursuit rotor) there is reason to believe that in concept formation problem solving and creative thinking bidirectional associations are the rule The expectations are logical but it is difficult at the time to find evidence for them So to a large extent such phenomena have been omitted in the above discussion although a paper written five years from now may include them Indeed they may be included by other more perceptive participants in this conference

Looking at possible integrative attempts in future years one may ask

what factors in the past have seemed to be responsible for the relative isolation of work in the various areas of learning It is the opinion of the present writer that there is one basic habit pattern possessed by most of us which tends to restrict us, this is a severe adherence to particular methods of investigation While this behavior may be quite understandable, and while standardized methods may have many merits, their possible evils cannot be gainsaid Standardization of methods in its most severe form may have produced or contributed to the very problem this conference is intending to resolve Some expansion on this point is obviously required

Learning theories, as developed in the animal learning laboratory, have never seemed to the writer to have much relevance to the behavior of a subject in learning a list of paired associates The emphasis upon the role of a pellet of food or a sip of water in the white rat s acquiring a response somehow never quite seemed to make contact with the human *S* learning to say VXK when the stimulus DOF was presented Furthermore, it does not seem to me that anyone else has been able to show these relationships in any substantial way Are we to accept a conclusion that we will have different principles of learning for different species? Most of us would not accept this any more than we would accept the idea that we will have fundamentally different principles for different forms of human learning

Our so-called standardized tasks are very contrived and over the years have been simplified and refined They are contrived to allow only a limited range of behavior to be exhibited and they are simplified to allow specific stimulus response relationships to be derived and to be repeatable from one laboratory to another Certain response measures, and these alone, are recorded The many virtues present in this system are quite manifest, the dangers less so The highly standardized situations allow the *S*, whether rat or man, to behave in only a particular way and the range of this behavior in one situation may not overlap the behavior allowed in another What may be said to be a basic phenomenon in one situation is not found in another because the nature of this other situation is not such as to allow the phenomenon to be produced

As noted, the points of contact between theories based on a rat in a Skinner Box and the behavior of an *S* before a memory drum are difficult to find Could this be because the two situations simply do not allow certain principles of learning to be operative in a maximal manner? For example, the writer once had an *S* conduct an experiment in verbal learning which was essentially a free-operant learning procedure (Underwood & Schulz, 1960 pp 273–278) The *S* was presented stimulus terms just as in the traditional paired associate method but he was to supply his own responses these being restricted to three letter words The *S* was required

to attach a particular response to a particular stimulus and the task was completed when this was consistent on two successive trials One might expect this to be an easy task, but the results showed that it was not The *S* apparently had no trouble giving responses but to get them to stick to particular stimuli was difficult A response might be given to a particular stimulus for two or three trials and then it would drop out entirely or migrate to another stimulus It is very possible that we might come much closer to realizing the role of reinforcement in verbal learning than we do in the standard situation if, when the *S* had given a response to the same stimulus, say, on two trials, the experimenter starting saying right' when the response was given to the stimulus on subsequent trials Perhaps if some such procedure were used we not only would facilitate the rate of learning but could also generate the operant conditioning curves associated with the rat in a Skinner Box

It was said earlier that standardized tasks tend to allow only a limited range of behavior to occur This may be illusory in that along with a persistent use of a given task goes the persistent use of a given response measure Behavior which does not fit this measure is ignored and even scorned, with only the more dramatic forms living in the folklore of the laboratory The investigator who is studying running speed in a straight alley as a function of Variable X is irritated when the rat lies down in the alley and goes to sleep In our laboratory, whenever an *S* fails to complete the task for any reason, the experimenter must note the facts of the case on the record sheet One such sheet in an experiment where difficult lists were involved showed the following notation About half way through the experiment the subject rose from his seat muttered To hell with it,' and walked out of the room Another, recently noted was Subject dismissed, insisted on talking to the *E* when he was supposed to be learning"

Perhaps we are not alert enough to such happenings Perhaps the investigator of the sleepy rat could, by modifying his situation somewhat produce an ideal device for studying sleep inducing stimuli Perhaps the *S* who stalked out of the experimental room rather than continue learning should give us a lead for the study of frustration and withdrawal behavior Perhaps the talkative *S* has provided us with an hypothesis on The memory drum as an instigator of social behavior While these are facetious and extreme, they may be taken as illustrations of the fact that our standard situations, from which we measure only a limited form of behavior, may be in part responsible for our inability to relate adequately the learning behavior in one situation to that in another

The standard task when used over the years becomes a vehicle for

which pitfalls are known and we are thus much more able to avoid errors of method We feel at home with such tasks But the task may become synooymous with the cooceptual notions of behavior we are tryiog to understand, aod this can be an evil As discussed earlier, mediation is being given considerable prominence in verbal learning as well as in other areas In paired associate learning the 'standard' rate of presentation is 2 2 sec, that is, 2 sec for the presentation of the stimulus and the anticipation of the response and 2 sec for the presentation of the stimulus and response terms together If mediation of any degree of complexity is coosidered and is asserted to have a reality in the associatioo processes of the *S*, it is quite clear that a 2 sec anticipation interval is simply insufficient time for a chain of associatioos to be 'run off" Yet, how many studies have increased the anticipation interval to allow for this? Very few, and yet we get studies which cooclude that if mediatioo does occur it has little or no effect on the rate of learning How can mediation be expected to influence learning when the situation is such as not to allow time for mediation to occur

If in the science of learning we have anything approaching 'breakthroughs' it may be because these occur when the investigator is able to throw off the shackles imposed by standardized procedures For example, the innovation of dropping out items as iotroduced by Rock (1957) has led to a great deal of research which is resulting in a much better understanding of item selectioo processes and intralist effects That Rock's theoretical conclusioos may be quite unjustified is quite beside the point, we know more about verbal learning than we did aod it is doubtful if our knowledge would have taken this direction had it oot been for Rock s drastic variatioo in method Breaking away from a standardized procedure is dangerous for the investigator is often likely to think that the behavior of the *S* in the new situatioo will be the same as it was in the standard situation And this is the point namely, that new situations may iotroduce new forms of behavior over and above those we expect and by just this process we may reach out to another previously unrelated area of iovestigation and provide a bridge between the two

At a strictly intellectual level it is doubtful if any serious disagreemeots will be met with the notion that we should not take our standardized methods too seriously Yet, we wilt have difficulty translating sucb ioto action Our thinking is so geared to the standard situations that to change is difficult Undoubtedly it is more esthetically pleasing to let what changes in method we bring be dictated by theoretical considerations, but it may be equally instructive to examine our standardized procedures and say, now what would be the implications if this were changed or that were changed In the discussion of verbal learning in this paper it was clear that

one of the processes of learning involved the extinction of old habits The investigator in a science is a learner—an *S*, some of his habits may also need periodic extinction

References

Alluisi, E. A, & Muller, P F, Jr. Verbal and motor responses to seven symbolic visual codes A study of S-R compatibility *J exp Psychol*, 1958, 55, 247–254

Barnes, Jean M, & Underwood, B J "Fate" of first list associations in transfer theory *J exp Psychol*, 1959, 58, 97–105

Bousfield, W A The occurrence of clustering in the recall of randomly arranged associates *J gen Psychol*, 1953, 49, 229–240

Bousfield, W A, Whitmarsh, G A, & Danick, J J Partial response identities in verbal generalization *Psychol Rep*, 1958, 4, 703–713

Briggs, G E Acquisition, extinction, and recovery functions in retroactive inhibition *J exp Psychol*, 1954, 47, 285–293

Cohen, B H, Bousfield, W A, & Whitmarsh G A *Cultural norms of verbal items in 43 categories* Tech Report No 22, ONR Contract Nonr 631(00), Univer of Connecticut, 1957

Deese, J Influence of inter item associative strength upon immediate free recall *Psychol Rep*, 1959, 5, 305–312

Dicken, C F Connotative meaning as a determinant of stimulus generalization *Psychol Monogr*, 1961, 75, Whole No 505

Dodwell, P C Coding and learning in shape discrimination *Psychol Rev*, 1961, 68, 373–382

Eckstrand, G A, & Wickens, D D Transfer of perceptual set *J exp Psychol*, 1954, 47, 274–278

Epstein, W The influence of syntactical structure on learning *Amer J Psychol*, 1961, 74, 80–85

Fitts, P M, & Seeger C M S R compatibility Spatial characteristics of stimulus and response codes *J exp Psychol*, 1953, 46, 199–210

Goss A E Verbal mediating responses and concept formation *Psychol Rev*, 1961, 68, 248–274

Haagen, C H Synonymity, vividness familiarity, and association value ratings for 400 pairs of common adjectives *J Psychol* 1949, 30, 185–200

Horowitz, L M Free recall and ordering of trigrams *J exp Psychol* 1961, 62, 51–57

Horton, D L, & Kjeldergaard P M An experimental analysis of associative factors in mediated generalization *Psychol Monogr*, 1961, 75, Whole No 515

Howes, D, & Osgood, C E On the combination of associative probabilities in linguistic contexts *Amer J Psychol*, 1954 67, 241–258

Karwoski, T The memory value of size *J exp Psychol* 1931, 14, 539–554

Matthews, R Recall as a function of number of classificatory categories *J exp Psychol*, 1954, 47, 241–247

Mednick, S A, & Freedman J L Stimulus generalization *Psychol Bull*, 1960, 57, 169–200

Melton, A W Learning In W S Monroe (Ed), *Encyclopedia of educational research* New York Macmillan, 1941, pp 667–686

Miller, G A Free recall of redundant strings of letters *J exp Psychol* 1958, 56, 484–491

Osgood, C E, Suci, G J, & Tannenbaum, P H *The measurement of meaning* Urbana Univer of Illinois Press, 1957

Peterson, L R, & Peterson, M J Short term retention of individual verbal items *J exp Psychol*, 1959, 58, 193–198

Postman, L The present status of interference theory In C N Cofer (Ed), *Verbal learning and verbal behavior* New York McGraw Hill, 1961, pp 152–178

Postman, L, & Riley, D A A critique of Kohler's theory of association *Psychol Rev*, 1957, 64, 61–72

Richardson, J The relationship of stimulus similarity and number of responses *J exp Psychol*, 1958, 56, 478–484

Riley, D A Rote learning as a function of distribution of practice and the complexity of the situation. *J exp Psychol*, 1952, 43, 88–95

Rock, I The role of repetition in associative learning *Amer J Psychol* 1957, 70, 186–193

Schulz R W, & Runquist, W N Learning and retention of paired adjectives as a function of percentage occurrence of response members *J exp Psychol*, 1960, 59, 409–413

Shepard, R N, Hovland, C I, & Jenkins, H M Learning and memorization of classifications *Psychol Monogr*, 1961, 75, Whole No 517

Twedt, H M, & Underwood, B J Mixed vs unmixed lists in transfer studies *J exp Psychol* 1959, 58, 111–116

Underwood, B J Studies of distributed practice IX Learning and retention of paired adjectives as a function of intralist similarity *J exp Psychol*, 1953, 45, 143–149

Underwood B J An evaluation of Gibson theory of verbal learning In C N Cofer (Ed), *Verbal learning and verbal behavior* New York McGraw Hill, 1961, pp 197–217

Underwood B J Stimulus selection in verbal learning In C N Cofer & B S Musgrave (Eds), *Verbal behavior and learning Problems and processes* New York McGraw Hill 1963

Underwood B J, Ham, M, & Ekstrand, B Cue selection in paired associate learning. *J exp Psychol*, 1962, 64, 405–409

Underwood, B J, & Richardson, J Studies of distributed practice XVII Interlist interference and the retention of paired consonant syllables *J exp Psychol*, 1957, 54, 274–279

Underwood B J, Runquist, W N, & Schulz, R W Response learning in paired associate lists as a function of intralist similarity *J exp Psychol*, 1959, 58, 70–78

Underwood B J & Schulz, R W *Meaningfulness and verbal learning* Philadelphia Lippincott 1960

Underwood B J, & Schulz R. W Studies of distributed practice XXI Effects of interference from language habits *J exp Psychol* 1961 62, 571–575

Waters, R W The law of acquaintance *J exp Psychol*, 1939, 24, 180–191

The Centrality of Verbal Learning

COMMENTS ON PROFESSOR UNDERWOOD'S PAPER

DELOS D WICKENS
Ohio State University

I fear that for me to discuss the specifics—the contents and empirical findings—of verbal learning following a paper on this topic by Underwood is like carrying a very poor grade of coal to Newcastle, so I shall let the specifics go and concern myself with the more general matters that he has raised The first of these deals with the very basic reasoning behind this conference, the question of the degree of communal sharing of principles of learning—or perhaps performance—among the multiplicity of tasks devised for man by nature and by himself The second of these is a query into Underwood s statement that ' the work in verbal learning—rote verbal learning—may stand squarely in the center of all human learning (p 52)

UNITY VS DIVERSITY OF LEARNING PROCESS

My attitude toward the first of these topics is predictable from my early academic history, for I was a student of J F Dashiell At this time—in the 1930 s—the multiplicity proponents offered three major varieties of learning the conditioned reflex, trial and error learning and insight In short, the war between Behaviorism and Gestalt was still hot Dashiell had published in 1935 a paper entitled A Survey and Synthesis of Learning Theories ' in which he carefully examined the possibilities of finding similarities rather than differences among these presumably different kinds of learning Part of his conclusion (1935, p 275) read as follows

> All such extreme characterizations then are caricatures and as these supposedly differentiating traits lose their sharpness of focus the three kinds of learning lose some of their disparaleness and contradiction and a description and interpretation of all learning phenomena in general seems a sensible possibility

The case he made for his position was a potent one and then and there I was conceptually imprinted I was led to hope for, as well as to believe in, the possibility of formulating certain general principles—or as they are now more often called, dimensions of psychological processes (Melton,

1941)—which contribute in varying degree to the performance on the multiplicity of learning situations with which we can be confronted.

It is perfectly obvious that the number of these tasks approaches the infinite and to assume that each one has its own laws would be to say that we cannot reduce the complexity of the behavioral universe. Were such the case I think we would not be scientists, for the aim of a science is not only to find order in things, but to simplify and to do with as few concepts as possible. This aim toward parsimony, as well perhaps as the desire for the aesthetic thrill which comes from reducing the complicated to the simple (as mathematics so nicely exemplifies), should prevent us from accepting too readily a position of multiplicity of learning processes. It should lead to the type of realistic dissatisfaction that pushes one to try other means and other formulations which are as effective in prediction but which achieve the same degree of precision with a greater economy of concept. Perhaps, of course, this may be impossible of attainment, but there is no reason to accept defeat without trying. Thus it seems to me, as obviously it also does to Underwood, that the search for continuity across all human learning if not within our grasp should always be the aim of our reach, and that we should be chary of accepting a conclusion of a multiplicity of independent learning situations, though still, of course, prepared to accept the multiplicity conclusion if a more parsimonious one fails to account for the data.

Since it is undoubtedly agreed that one need not generate an independent set of principles for each learning situation, and that we should seek for simplicity, the problem arises of finding criteria which convince us that one class of learning situation requires a set of principles which are independent of some other, or at least independent in one or two respects. I believe that the valid approach is one that demonstrates that the $R = f(s)$ laws (or the antecedent conditions, and the nature of their relationship to the psychological construct) are different in one class of situation from what they are in another. This approach was used by Tolman (1949) in his article "There Is More Than One Kind of Learning" and some examples from this paper will illustrate the meaning of the statement. In speaking of the deacquisition of a positive cathexis, a concept similar to secondary reinforcement, he says that he doubts that it is ever, or at least much, weakened by the mere passage of time and lack of exercise. So also is this true of negative cathexes. But with respect to Field Expectancies, he states "true forgetting (i.e., weakening as a result of the mere passage of time) *does* take place . . . this forgetting obviously obeys the sorts of laws which the Gestalt psychologists have uncovered and not the old simple associationistic ones" (Tolman, 1949, p. 152). It is clear, therefore, that in Tolman's thinking the procedural operations which are required for the

decrement in strength of one of these processes are different from those of the other Another example in Tolman s system of the disparateness of antecedent conditions is found in the acquisition of field expectancies and positive cathexes For the former, drive reduction is irrelevant and the reinforcing agent plays a relatively minor role, but with respect to the acquisition of a cathexis, he states that he believed, like Hull, in the 'efficacy of reinforcement or need reduction" (1949, p 146) In other words, the antecedent conditions related to variation of one kind of learning differed from those of the other

Spence (1956) has used the same approach in his differentiation of the instrumental avoidance situation of the type represented by the eyelid conditioning from an approach learning represented by an animal scurrying down a runway for its daily pellet Spence is quite specific, the antecedent of H (the intervening variable leading to the running response) in the latter case is a function only of N and not of incentive magnitude, in the former it—the intervening variable H—is a function of N and also of magnitude of the UCS This conclusion leads him to describe the excitatory component of behavior, E, as being a function of two intervening variables, H and D, insofar as classical aversive conditioning is concerned, while the excitatory component of runway behavior requires, for him, three intervening variables, H, D, and K

Whether one agrees with the ultimate conclusion which these men draw about kinds of learning is not a matter of concern at present, the point is that they are using what seems to me the only acceptable criterion for concluding whether or not the learning in different situations or kinds of responses should be called different kinds of learning Without such a rigorous criterion as this one is likely to be misled by superficials and stray from the path of scientific righteousness

Since I must assume that our intentions are to seek parsimony and a sort of classical purity in theory construction I should speculate on the blandishments and deceits which lead us down the garden path of promiscuous theorizing Underwood has mentioned some of these, he stressed for example, the blending and binding effect of standardized tasks, and I do not think there are many who gainsay him May I add and/or emphasize a few more

There is to begin with the inadequacy of knowledge, an inadequacy which may be due to the fact that some field of investigation is so new that we know practically nothing about its empirical parameters and, hence are inclined to make statements about its unique characteristics which we would not do if the knowledge of behavior within that field were deeper A recent instance of this sort is to be found in the field of short term memory In their excellent study which essentially pioneered a new method

for the investigation of short term memory, the Petersons (1959) could find no evidence for proactive inhibition as a variable accounting for their obtained decline in retention This negative finding implied that short term memory and long term memory were not a function of the same variables and hence could be considered as fundamentally different from each other Later research by Keppel and Underwood (1962) has demonstrated proactive inhibition is indeed a powerful variable in the Petersons' short term memory paradigm, but particular aspects of the Petersons' procedure had masked its manifestation Thus, as additional data became available, an earlier distinction between types of psychological phenomenon did not seem to be required

I need touch but shortly on my next point since Underwood also has mentioned it, for it has to do with the nature of the task We ordinarily develop, invent, or contrive tasks so as to emphasize some aspect of behavior in which we are interested In a given environment—in this case the experimental task—there is only a limited number of behavioral characteristics which can be evoked or measured and so the knowledge derived by the experimenter must in turn be limited The ordinary classical conditioning situation tells us little about stimulus selection, for not much of a problem in stimulus selection is given to the *S* as he sits in a darkened, silenced room with earphones on his head awaiting the onset of the tone which precedes the shock Nor can the Skinner Box research readily deny a relationship between stimulus and response for instrumental learning in general when *E* has only a general control over *S*'s stimulus environment, the control being established only by putting *S* in the box Interestingly enough, when the Skinnerian begins to work with S^D s and S^Δ's, he begins to talk about the stimulus controlling *S*'s behavior, which is suspiciously similar to the Pavlovians' statement that the CS elicits the CR The fact that two tasks may differ in the sense that a particular psychological dimension is present in one and not in the other may mean that a theoretical account of the behavior in one task will include terms which are not required for the theoretical description of the other A condition of this sort, however, does not demand the conclusion that the two tasks represent basically different kinds of learning processes

A somewhat less obvious difficulty arises from the characteristics that the *S*s bring to the task and the way these characteristics interact with performance Harlow (1949) has strikingly demonstrated the difference between the performance of monkeys after only a few groups of discrimination problems and after many If our study of discrimination learning were limited to only the naive or to the sophisticated animals we might develop quite different theories to account for the obtained data, depending upon the class of *S* used There are, of course, other kinds of important subject

variables, and these may interact with the task in such a way as to imply that the task laws differ when in fact this interpretation may not be necessary Thus it has been found that the greater resistance to extinction under partial reinforcement as opposed to continuous holds if the rewards are viewed as being somewhat fortuitously given but are reversed if they are considered to be a function of skill (Rotter Liverant & Crowne 1961) The authors are able to interpret these findings not by assuming that the laws of learning differ for a skilled and a chance task but rather on the basis of what *S* considers the reinforcement to be in these two types of task

In summary I have taken the position that although we may be required eventually to conclude that there are a diversity of learning processes it is not a conclusion which should be accepted without careful scrutiny The conclusion can be justified only if it can be demonstrated that the antecedent conditions of which performance is a function in one class of situation are essentially different from those of the other class But even before this conclusion is drawn one must be assured that the generalization about learning processes is not the outgrowth of the limited scope of the tasks under consideration nor of limited knowledge of all the variables influencing these tasks

The Centrality of Verbal Learning

Now I would like to comment upon Underwood's statements that the work in verbal learning—rote verbal learning—may stand squarely in the center of all human learning (p 52) and that it is shooting out phenomenon and theories which are touching all areas of human learning from simple conditioning to the study of the thought processes (p 52)

As a species man is unique in the extent of his use of the verbal response system and most of our achievements in changing our environment and even our very selves from the uncivilized to the civilized must be attributed to our competence and precision in usage of this response system Certainly then the study of verbal learning is the study of the kind of behavior which most clearly differentiates the human from other forms of animals and it is the study of a kind of behavior which pervades or can be made to probe into nearly or perhaps all of our activities In this sense then verbal learning stands as the favored candidate for centrality in human learning But Underwood has gone a step further he specifies rote verbal learning not just verbal learning and it is probably the rote characteristic of these studies which has led its detracters to accuse this work of being—as Underwood mentions—dull narrow sterile and the like

My effort to evaluate Underwood's statement will begin by considering

what categories we need employ to contain *all of human learning* The categories that I am going to suggest are not based upon response systems, but rather upon activities which are functionally important in determining the organism's commerce with his environment There is no pretense of originality and they are included in some form or another in Tolman's list (1949) but I do not wish to imply that they necessarily represent different kinds of learning They are without question, however, different psychological processes which contribute to overall performance and which are modifiable or learnable

Four categories will do for my purposes (1) The first I will call *motivational learning* It is the acquisition of predispositions toward particular things and states of affairs It includes predispositions to approach and predispositions to avoid (2) The second consists of defining or—to borrow a word from Dewey—*constituting the stimuli qua stimuli* It is a learning by the organism of how the complexities of the impinging physical energies may be ordered, simplified, or differentiated one from another (3) The third is *response learning*, the acquisition and ordering of the motor coordinations which constitute the action of the organism on its environment (4) The fourth, of course, is *instrumental learning* It is a form of particularizing a relationship between a stimulus, a response, and a motive That is, it is a learning which is valid for the organism only if the end result of responding changes in a favorable way the relationship between the consequence of responding and the previously activated personalized states of the organism, i e , his current motivation It is a kind of learning which makes no sense unless the three—stimulus, response, and motive change—are considered together Stimulus learning can be conceptually isolated, in the sense that possibly any stimulus can be at some time a cue for any response and hence may be studied for itself, responses also have similar fickle characteristics—the same responses, or class of responses, may be performed in the presence of a wide variety of stimuli and even motives The motives in turn may find means for appeasement through a variety of responses The instrumental learning situation, however, contains an immediate and particularistic relationship between response and concurrent demands of the organism It is the study of factors influencing this relationship which concerns us in instrumental learning

Verbal Learning and Motivational Learning

But now to consider rote verbal learning in light of these kinds of activities Underwood has said that Learning theories, as developed in the animal learning laboratory have never seemed to have much relevance to the behavior of a subject in learning a list of paired associates The emphasis upon the role of a pellet of food or a sip of water in the white

rat s acquiring a response somehow never quite seemed to make contact with the human subject learning to say VXK when the stimulus DOF was presented (p 74) The lack of contact arises I believe not from a difference in organisms, or even primarily from differences in methodology—though this must contribute to some of it—but it arises primarily because the rote verbal learning program is seldom if ever directed toward an exploration of motivational learning

Typically, the verbal learning researcher attempts to hold motivational variables constant while he manipulates such variables as item similarity or presentation rate It is true of course that variations in motivational states do serve as the independent variables of verbal learning studies as when one studies the effects of different magnitude of reward (Miller & Estes 1961) or level of anxiety (Spence, Farber & McFann 1956) on task performance Investigations of this sort are not however, concerned with the *acquisition* of motivational tendencies and they are not therefore what I would classify as studies investigating motivational learning In contrast rat research has seemed to have moved more and more into the area of motivational learning and away from studies of instrumental activities such as occur in complex mazes discrimination boxes and delayed reaction equipment I suspect that Underwood s observation about the lack of contact between the significance of the rat s food pellet to the college sophomore s verbalization of a CCC arises to a large degree from the heavy emphasis on motivational variables and motivational learning among the animal studies contrasted with the relative lack of concern for this problem in the field of verbal learning

Parenthetically, and from the point of view of my bias toward communality of principles in different types of learning situations I think it is significant to note that although rat research and human verbal learning may not communicate with each other they do have a common interpreter —the field of conditioning—for both areas borrow generously of its concepts

It is in my other three areas of learning and in their interaction that rote verbal learning can make its bid for a central position in human learning and where theoretical formulations of potential significance to other tasks may be developed May I consider them briefly one at a time

Verbal Learning and Stimulus Selection and Differentiation

This is a problem which has entered the ken of the verbal learning theorists only lately, though it most certainly was a problem for *S*s who were learning the lists even back in G E Muller s day I suppose the lack of research on this problem arose primarily because other topics were given higher priorities and also because the researcher tried to eliminate need for

stimulus selection as an aspect of the overall learning Usually the experimenter is quite specific in his statements to the *S* as to what the class of stimuli and responses in the particular task are, and so there would seem to be little need for stimulus constituting or coding on the *S*'s part The *S*, in short, should use the stimulus in the manner in which the *E* makes it readily available to him However, as Underwood has remarked elsewhere, ' the college sophomore is a perfect confirmation of the law of least effort (1963, p 35) and he will reform or reconstitute the stimulus to suit his own tastes The recognition of the fact that processes of this sort do occur in the relatively simple verbal learning situation and that the nominal and the functional stimulus are not always identical establishes a common ground between a class of problems in the field of perception and the field of learning The potential advantage of this for the development of psychological theory is that the verbal learning theorist is in a position to study the process as it interacts with response learning as well as S—R formation He could also, by complicating the situation, move by gradual steps into the area of concept formation and problem solving

Verbal Learning and Response Integration

Somewhat the same can be said on the response integration side of the verbal learning situation Its independent contribution to overall learning as well as that of the stimulus has been abundantly evidenced in the work of Underwood and Schulz (1960) Although I do not believe it is exactly like the response integration in the tracking situation, it shares enough in common with it to promise cross fertilization

Verbal Learning and Instrumental Learning

I would classify what is called the associative or hook up phase of verbal learning under the rubric of instrumental learning Instrumental learning I had described as an instance where stimulus, response, and motive changes are considered together I assume that ordinarily the *S* obtains some kind of motive satisfaction when he verbalizes a response and then has that response show up immediately on the memory drum Thus giving *particular* responses to *particular* stimuli is instrumental in obtaining motive satisfaction for most of these *Ss* It is obvious that this final accomplishment of giving the proper response to each stimulus is a complicated function of stimulus learning response learning and then of an S—R associative or hook up phenomenon That each of these occurs while learning a list of paired associates has recently been most cleverly demonstrated by McGuire (1961) The verbal learning situation seems to me to be a

uniquely effective one for studying the interactions of these three kinds of learning activities.

Summary and Conclusion

The primary concern of this discussion has been with evaluation of Underwood's statement that work in rote verbal learning is central to all human learning. Working under the assumption that essentially all of human learning may be conceived of as falling into the four categories of (1) motivational learning, (2) stimulus constituting, (3) response learning, and (4) instrumental learning, the contribution of verbal learning studies to each of these categories was discussed.

It was concluded that studies dealing with motivational learning are negligible in the field of verbal learning, and that such studies are far more characteristic of the work with lower animals It was suggested that verbal learning situations seem to be especially useful for investigating interrelationships between stimulus learning, response learning, and the final product of performance—stimulus-response learning.

References

Dashiell, J F. A survey and synthesis of learning theories *Psychol Bull*, 1935, 32, 261–275.

Harlow, H. F. The formation of learning sets *Psychol Rev*, 1949, 56, 51–65

Keppel, G, & Underwood, B. J, Proactive inhibition in short-term retention of single items *J. verb. Learn. verb Behav.*, 1962, 1, 153–161

McGuire, W. J. A multiprocess model for paired associate learning *J exp Psychol*, 1961, 62, 335–347.

Melton, A. W. Learning In W. S Monroe (Ed), *Encyclopedia of educational research* Macmillan, 1941, pp 667–686

Miller, L. B, & Estes, B W. Monetary reward and motivation in discrimination learning *J. exp Psychol*, 1961, 61, 501–504

Peterson, L R, & Peterson, M. J Short term retention of individual verbal items *J. exp Psychol*, 1959, 58, 193–198

Rotter, J. B, Liverant, S, & Crowne, D P. The growth and extinction of expectancies in chance controlled and skilled tasks *J Psychol*, 1961, 52, 161–177.

Spence, K W. *Behavior theory and conditioning* New Haven Yale University Press, 1956

Spence, K W., Farber, I. E, & McFann, H H The relationship of anxiety (Drive) level to performance in competitional and non competitional paired associates learning *J exp Psychol*, 1956, 52, 296–310

Tolman, E C. There is more than one kind of learning *Psychol Rev*, 1949, 56, 144–155

Underwood, B J Stimulus selection in verbal learning In C. N Cofer & B S Musgrave (Eds), *Verbal behavior and learning Problems and processes* New York. McGraw Hill, 1963, pp 33–48.

Underwood, B J, & Schulz, R W. *Meaningfulness and verbal learning* Chicago Lippincott, 1960

Probability Learning[1]

WILLIAM K. ESTES
Stanford University

Like many other varieties of learning, probability learning is easier to identify than to define The various familiar types of learning are the result not of any rational scheme of classification, but simply of the clustering of research activity around points of communality in problems and procedures By the usual criterion, human probability learning amply qualifies as a "type" Although the bibliography assembled for this article is by no means complete, there are represented more than 80 separate experimental studies, many including numerous subexperiments, that have been published during the last 10 to 12 years, all primarily concerned with probability learning in human subjects (*Ss*)

The basic paradigm for the probability learning experiment has become almost as familiar as that for classical conditioning On each of a series of trials, the *S* makes a choice from an experimenter defined set of alternative responses, usually though not necessarily verbal, then receives from the experimenter a signal indicating whether the choice was correct In the most common experimental design, the simple noncontingent case, each response has some fixed probability of being reinforced (indicated as correct) on any trial, regardless of the *S*'s present or past choices In the oldest variant (Humphreys, 1939, Grant, Hake, & Hornseth, 1951) the *S*'s assigned task is predicting which of two alternative events (e g light on or light off) will occur, and at the end of each trial exactly one of these events does occur

Although one of the main objectives of this paper is to attain some perspective as to ways in which probability learning might fit into a systematic treatment of human learning, I shall not linger at the outset over criteria of classification I suggest that to be of any lasting scientific value a taxonomy of human learning will have to be based upon theoretically significant distinctions and communalities consequently progress toward a useful taxonomy may be facilitated if we examine the phenomena of probability learning and phenotypically similar varieties of learning at several levels of analysis A natural order of procedure will be first to review the materials, i e , methods, empirical findings and descriptive

[1]Preparation of this review was supported in part by Contract Nonr 908(16) between the Office of Naval Research and Indiana University Reproduction in whole or in part is permitted for any purpose of the United States Government

models, then to undertake comparisons with some of the categories of learning for which similar reviews have been provided in other papers of this symposium

Methodological Approaches

Problem-Oriented Investigations

To some extent in most studies of probability learning, and almost exclusively in a few, the investigator is simply interested in gaining an understanding of this behavioral situation as an end in itself The natural outcome of this approach is an accumulation of factual material and the growth of what may be termed "situational theory" (i e, interpretation of the phenomena in terms of concepts and assumptions which may or may not have some generality taken individually but which in combination form a theory for this particular type of situation) Studies by Gardner (1957, 1958), Jones (1961), Rubinstein (1959), and, perhaps somewhat less clearly, Goodnow and her associates (Goodnow, 1955, Goodnow & Pettigrew, 1955, Goodnow & Postman, 1955, Goodnow, Rubinstein & Lubin, 1960) seem to exemplify this problem oriented approach

Theory Oriented Investigations

At the other extreme are studies which simply take the probability learning situation as a convenient testing ground for general theories (e g, Estes & Straughan, 1954, Kalman, 1961, LaBerge, 1959a, Suppes & Atkinson, 1960) Recalling that the simple predictive ("verbal conditioning") situation was originally devised by Humphreys (1939) as a verbal analogue to Pavlovian conditioning, one should not, perhaps, be surprised that the theories involved have for the most part been learning or conditioning theories Recently, however, decision models have been applied, in some cases for the specific purpose of testing predictions drawn from game or decision theory against predictions derived from learning theory

Although these two approaches often tend to shade into one another in practice, it is worthwhile recognizing the clear distinction between them in principle, if only because one may thus avoid considerable fruitless controversy Testing a general law or model in a particular experiment is not tantamount to proposing it as a situational theory capable of accounting for all aspects of behavior in the experimental situation One may apply, say, the laws of motion and conservation of energy to the behavior of an automobile (in fact one is well advised to do so), but with no idea of deciding that the laws are "wrong" if they do not suffice to predict every thing the vehicle may be found to do I shall take the position that a general theory is supported by experimental tests if (*a*) under specifiable

conditions it leads to correct predictions, and if (*b*) analyses of conditions under which the theory does and does not hold show that they fit some orderly pattern and thus further our understanding of the phenomena concerned The extent to which certain general theories are supported, in this sense, by the data of probability learning experiments will be one of the principal points at issue in the following sections

Use of Group vs Individual Data

Nearly all of our knowledge about probability learning is based on statistical analyses of group data The familiar negatively accelerated learning curve is a curve of group mean proportions, asymptotic probability matching is a correspondence between terminal mean response proportions and proportions of reinforcing events, and so on In this respect, the situation in the study of probability learning is much like that in most other areas of human learning But in the case of probability learning one more often hears questions raised as to what the individual *S*s are doing Do their learning functions follow negatively accelerated curves? Do their response proportions individually match reinforcement probabilities? The answer, of course, is that the individual learning functions are quite irregular and generally reveal no simple properties at all to the naked eye

Consider, for example, the individual learning functions shown in Table 1, these represent block by block proportions of the more frequently reinforced response for 8 *S*s [the first 2 *S*s from each of four subgroups in the 85 15 condition of the Estes and Straughan (1954) study] It is apparent at a glance, firstly, that there are sizable individual differences at all stages of learning and, secondly, that none of the *S*s exhibit a uniform

TABLE 1

RESPONSE PROPORTIONS BY 20-TRIAL BLOCKS FOR INDIVIDUAL *S*s RUN UNDER AN 85 15 SCHEDULE

	Block				
Subject	1	2	3	4	5
1	65	85	85	85	95
2	60	80	85	85	60
3	55	90	75	90	90
4	55	90	90	1 00	1 00
5	80	95	85	90	85
6	45	60	45	70	70
7	55	90	90	95	90
8	60	90	95	75	85

increase in response probability toward the theoretical asymptote The mean proportion for the group on the last block (84) approximates the matching value, but only 2 of the 8 *S*s fall close to the mean From inspection of these individual functions, one might believe it quite unlikely that all could have been generated by the same basic learning process

However, impressions gained from observations of individual cases without adequate control comparisons are notoriously fallible To provide a basis for comparision, I have assembled in Table 2 some individual learn

TABLE 2

RESPONSE PROPORTIONS BY 20-TRIAL BLOCKS FOR SIMULATED DATA GENERATED BY COMPUTER FOR STATISTICAL '*S*s' LEARNING UNDER A 75 25 SCHEDULE

Subject	Block 1	2	3	4	5
1	60	60	75	80	80
2	80	70	70	90	1 00
3	65	85	80	80	65
4	70	45	65	65	70
5	65	60	85	80	95
6	35	65	55	40	80
7	80	85	50	75	80
8	55	85	75	85	90

ing functions obtained from artificial *S*s by means of a computer The computer was programmed to generate simulated data for '*S*s' who learn in accord with the linear model of statistical learning theory The records shown in Table 2 are not selected cases, but merely the first eight cases of a large group run for another purpose The absolute values for the simulated *S*s are not comparable to those for the real *S*s, since the former were obtained with reinforcement probabilities of 75 25 rather than 85 15 However, it is interesting to note that, except for the overall level, the patterns of variation in the two sets of data are strikingly similar Individual differences are quite comparable in the two sets of data at all stages of learning, if anything the range of individual differences is a bit larger for the simulated *S*s—although they are known to have identical learning parameters The artificial data give us an idea of how much variability to expect both among individuals at any stage and over trial blocks for any individual simply because of the probabilistic character of the choice

behavior, even when there are no individual differences in mode or rate of learning

The learning functions of Table 1 include examples of a number of the idiosyncrasies which, in various published studies, have been taken as evidence supporting hypotheses concerning individual differences in decision rules and the like Thus, S No 2 exhibits a sudden regression toward chance in the last block, as though he had finally decided to abandon a strategy which had not proved fully satisfactory (as does simulated S No 3 in Table 2), S No 4 appears to have found a 'maximizing' strategy (as does simulated S No 2), S No 6 shows a long initial period of no improvement, as though he were trying and rejecting various irrelevant hypotheses rather than learning in any orderly manner as a function of reinforcements (and the same is true of simulated S No 6) Contemplation of these parallels between real and simulated data might if nothing else make one somewhat wary of schemes involving the classification of learners into types on the basis of their protocols

With respect to response proportions in the terminal trial blocks, one certainly would not be led from inspection of the individual protocols to say that probability matching characterizes all of the simulated Ss All of the simulated data are generated by a learning process which has the property that individual response proportions must vary randomly around a negatively accelerated function which goes asymptotically to the probability matching level But if we were given a quantity of similar data 'blind' and asked to deduce the nature of the generating process it is plain that we would be unlikely to discover this property except by appropriate statistical analyses including the averaging of learning functions for groups of Ss

The main point I wish to make by means of this example is that, for a given research area, arguments over the values and dangers of averaging can be brought down from the level of philosophical disputation to that of a problem for objective investigation In the particular case of probability learning, it seems clear that, whereas on the one hand incautious inferences from averaged data may lead us to attribute to individuals relationships which exist only in groups, on the other, failure to make judicious use of statistical analysis may result in our failing to discover relationships which do characterize the individual but which become apparent only in appropriately averaged data Having no sure safeguard against any one type of error, I can see no reasonable alternative to depending on the self corrective properties of a versatile approach including analyses in terms of statistical and other mathematical models as well as the intensive observation of individual cases

Some Facts about Probability Learning

Before entering on the uneasy ground of systematics, it may be useful to assemble some supplies by way of factual material A summary of empirical relationships will serve also to organize some comments about the roles of variation in a number of important parameters of the experimental situation It should be mentioned at the outset, however, that to hold this paper to manageable dimensions, it has seemed advisable to limit consideration to studies of the simple predictive situation with discrete choices and informational feedback As the comparative analysis begun in the present study is continued and amplified, a number of variations on the basic paradigm will merit attention these include studies of situations permitting continuous response measures (e g , Anderson & Whalen, 1960, Brunswik & Herma, 1951, Lotsof, 1959, Suppes & Frankmann, 1961), studies of effects of relative frequency of reinforcement in paired associate or other forms of verbal learning (e g , Binder & Feldman, 1960, Osgood & Anderson, 1957, Peterson, 1956, Solley & Messick, 1957, Voss, Thompson, & Keegan, 1959), studies of various reinforcement contingencies arising in social interactions (e g, Flood, 1954a, 1954b, Hays & Bush, 1954, Kanareff & Lanzetta, 1960, Neimark & Rosenberg 1959), and comparisons of probability learning in adult human *Ss* with probability learning in children or animal *Ss* (e g , Overall & Brown, 1959, Stevenson & Zigler, 1958)

Probability Matching in the Noncontingent Case

By the time the first half dozen studies of simple, two-choice predictive behavior (summarized in Estes, 1961b) had shown close approximations to probability matching with a degree of replicability quite unusual for quantitative findings in the area of human learning, it was only natural that investigators as a group should feel themselves challenged to push this phenomenon to its limits and see if the simple pattern could not somehow be made to break down I do not know whether anyone ever claimed probability matching to be a ubiquitous empirical generalization, but in any event several years of vigorous limit pushing on the part of a large number of investigators have served to delineate reasonably clear bounds on its region of applicability

Under the conditions of the early experiments, i e , a two-choice situation with reinforcement provided by a signal indicating the correct prediction on each trial curves of response proportions vs trial blocks approximate matching values closely for several hundred trials [for an especially clear case obtained with a very large sample of *Ss*, see Neimark and Shuford (1959)] In some experiments falling in this category a slight 'over

shooting" begins to develop after about 250 trials For example, groups run by Gardner (1957) with reinforcement probabilities of .60 and .70 yielded mean response proportions of .62 and .72, respectively, over the final portion of a 450-trial series

As various parameters of the experimental situation have varied more or less randomly over a considerable number of investigations, enough information has accumulated to permit some tentative generalizations about the conditions under which different types of asymptotic behavior arise

Probability matching tends to occur when the experimental task and instructions are such as to lead the *S* simply to express his expectation on each trial (Estes & Straughan, 1954, Friedman, Burke, Cole, Estes, Keller, & Millward, 1963) or when they emphasize the desirability of attempting to be correct on *every* trial, as in a simulated psychophysical situation (Estes & Johns, 1958) or problem solving situation (Goodnow, 1955, Goodnow & Postman, 1955) "Overshooting" of the matching value tends to occur when instructions indicate, directly or indirectly, that the *S* is dealing with a random sequence of events (Edwards, 1961, Morse & Runquist 1960, Rubinstein, 1959) or when they emphasize the desirability of maximizing successes over blocks of trials (Das, 1961)

When differential rewards (or punishments) are given for correct and incorrect responses, the probability of the more often rewarded alternative is frequently observed to go above the matching value, evidently approaching an asymptote which is directly related to the reward differential (Edwards, 1956, Siegel & Goldstein, 1959, Suppes & Atkinson, 1960) It would be premature, however, to conclude that learning under differential reward contingencies differs in any fundamental respect from learning in situations not involving reward In fact, it has been possible to show that terminal response proportions in a number of these reward experiments can be accounted for on the assumption that the behavior is compounded of two processes, each of which satisfies probability matching even though the observed choices do not (Estes 1962b)

Studies of noncontingent situations with more than two choices have yielded results in agreement with those of comparable two choice situations up to a point, but they have also turned up some puzzling discrepancies One of the first reported three choice experiments was conducted by Neimark (1956) under conditions otherwise analogous to those of several two choice experiments run more or less concurrently (e g , Burke, Estes, & Hellyer, 1954, Estes & Straughan, 1954) Neimark's terminal response proportions approximated the probability matching level very closely Neimark's experiment involved only 100 trials, but the possibility that she merely happened to terminate her series at the point when the empirical curves were crossing the matching level, as suggested by Gardner (1957),

can apparently be discounted in the light of later results obtained in the same laboratory with longer series To give one example, groups of 48 *Ss* were run under schedules of 60 20 20, 50 30 20, and 40 40 20 in an unpublished study by C J Burke, R Ginsberg, and the writer Mean proportions of the most frequently reinforced response over the last 40 trials of a 240 trial series were 593, 480, and 382, respectively It seems clear that under the conditions of these experiments the learning functions are quite well described by theoretical curves which have their asymptotes at the probability matching values

At the same time, it has become evident that with sufficiently long trial series, the proportion of choices of the preferred alternative tends eventually to rise above the matching value (Cotton & Rechtschaffen, 1958, Gardner, 1957, 1958, McCormack, 1959), this overshooting becomes manifest at about the same point as in comparable two-choice experiments, but appears to be somewhat more pronounced by the end of a 400- to 500 trial series The experiments cited above have utilized too few event sequences to permit reasonable estimates of mean response proportions at any stage of learning However, a replication of one of Gardner's conditions (the 70 15 15 condition from the 1957 study) run in the Indiana laboratory (Ginsberg 1959) with larger samples of both *Ss* ($N = 48$) and sequences ($N = 12$) yielded apparently stable terminal probabilities for the 70 alternative in the neighborhood of 75 (mean proportions being 756, 755, 759, 744, and 751 for the last five 20 trial blocks of a 460 trial series) More information is needed regarding the extent to which the asymptotic deviations may be peculiar to the type of reinforcement schedule used in all of these studies (i e, with one alternative having a much higher reinforcement probability than the others)

Problems in the Interpretation of 'Asymptotic Behavior'

In order to establish generalizations about asymptotic behavior, it is necessary to be able to ascertain when behavior is asymptotic But accomplishing this task, even for the simplest case of two-choice behavior in the noncontingent situation has proved inordinately difficult In the earliest studies, it was noted that learning curves typically level off after a hundred trials or so and consequently response proportions over the terminal portion of an 80- to 150 trial series were taken as estimates of asymptotes In response to doubts on the part of some investigators as to whether the curves were really asymptotic,' longer and longer series have been used running to 300 450 1000, and 1200 trials But, unfortunately, it becomes ever more doubtful that any satisfactory solution can be reached

by this frontal approach Relatively minor, although difficult enough, is the statistical problem of securing general agreement on a criterion for stationarity of response proportions over trial blocks, little has been accomplished with this problem, but at least in principle it can be solved

Deeper problems arise from the fact that, when we speak of asymptotic behavior, we mean "asymptotic under constant conditions" In trying to establish suitable conditions, we find ourselves at cross purposes If we run *S*s in long continuous sessions, it is probably impossible to maintain reasonable constancy with respect to motivation, alertness fatigue or boredom, and the like If we follow the example of the psychophysicists and attempt to mitigate these factors by using relatively short, well spaced sessions, we run into the difficulty of keeping the *S*s from thinking or talking about the experimental task between sessions, or in other ways obtaining information beyond that supplied by the experimental sequence of events The experimental procedures used in probability learning represent compromises between these conflicting considerations

Even when a learning curve levels off at some value of response probability intermediate between zero and unity, there are reasons to believe that under many circumstances this apparently asymptotic response level will be inherently unstable In nearly all probability learning situations, the nature of the task and the instructions indicate or imply that the *S* is expected to improve the accuracy of his predictions or judgments over a series of trials Consequently, if under a given set of conditions the *S*'s probability of being correct has remained constant over a substantial number of trials, it seems not unlikely that *S* will eventually recognize the absence of further 'improvement' and will vary his behavior We know that under suitable instructions, the *S* can respond to changes in the density of "correct" or "incorrect" signals over longer blocks than single trials as reinforcing events Consequently, we must admit the possibility that even under instructions which in effect define the individual trial outcome as the reinforcing event, a prolonged period of no improvement might instigate *S*s to shift their behavior in this direction If this analysis is correct it implies that theories couched in terms of choices and reinforcing outcomes of single trials cannot provide an entirely adequate account of long term changes in response probability, except perhaps under rather special conditions It does not rule out the possibility that such theories can gain support from predictions concerning curves of probability learning (and it by no means guarantees that more adequate theories will be easy to come by)

Sequential Properties of Behavior in the Noncontingent Case

Although it is apparently unrealistic to expect more than moderate success in predicting "asymptotic" response levels one may still hope that a

theory of probability learning can generate accurate predictions of response probability on a given trial of a series, given appropriate information about the immediately preceding sequence of events. Results in this category which have attracted the most widespread interest are those having to do with "recency curves."

Extant versions of reinforcement theory, as they have been applied to the two-choice, probability learning experiment, predict positive recency curves. That is, at any point in a series of trials, the greater the number of consecutive, immediately preceding occurrences of a given event, the more likely the *S* should be to predict that event on the next trial. Evidence bearing on this prediction was first reported by Jarvik (1951), and proved sharply negative. Rather than following the predicted function, *S*s tended to respond in terms of the "gambler's fallacy," that is, to be more likely to predict an event the longer it had been since the last preceding occurrence of the event. This "negative recency" effect was also found by a number of subsequent investigators (e.g., Anderson, 1960, Feldman, 1959, Nicks, 1959), the characteristic recency curve showing an increase in probability of a response after one or two reinforcements, then a decrease as the run of consecutive reinforcements continues.

A satisfactory interpretation of this phenomenon can hardly be expected until more information is available concerning its course of development and stability. The suggestion has been put forward (Estes, 1962a) that the negative recency function results largely from response tendencies the *S*s bring with them to the experiment, via generalization from other situations, but which extinguish with experience in the experimental situation. Some support for this interpretation has been provided by recent studies which have utilized sufficiently long series to provide a reasonable opportunity for the posited extinction to occur (Anderson, 1960, Edwards, 1961, Friedman et al., 1963). If previous analyses of the probability learning situation are extended, as in Restle's (1961, Ch. 6) treatment of choice behavior, to allow for the *S*s' responding to runs of reinforcing events of different lengths as differential cues, a reinforcement theory could accommodate a transitory negative recency effect. A bowed recency curve would be expected to develop early in a series simply because learning would proceed more rapidly with respect to cues from the shorter, and therefore more frequently occurring event runs. Persistence of a negative recency effect in asymptotic data would be compatible with Restle's model but very difficult to account for in terms of reinforcement principles.[2]

[2] Provided that the experimental procedure avoids artifacts such as those that arise from imposing constraints on the relative frequencies of reinforcing events within short trial blocks which would lead to differential reinforcement of negative recency behavior.

Another active line of investigation has been concerned with attempts to show that the effective stimulus for the *S* at any point in a probability learning series is the immediately preceding run of similar reinforcing events There is ample evidence that *S*s respond differentially to the stimulus aftereffects of the immediately preceding trial, this is seen most directly in experiments which arrange different probabilities for a given reinforcing event depending on which event occurred on the preceding trial (Hake & Hyman, 1953, Engler, 1958, Anderson, 1960, Shelly, 1960a, 1960b) Attempts to show that *S*s discriminate the differential stimulation associated with different lengths of runs of a given event have not yielded entirely conclusive results, partly because the investigators have not used available analytic techniques to full advantage for the purpose of partialling out the reinforcing effects of a given event run from its effects, if any, as a discriminative stimulus (see, e g , Feldman, 1959, Goodnow, Rubinstein, & Lubin 1960, Nicks, 1959) One would surely expect that adult human subjects could, up to some limit, respond differentially to stimulation from different lengths of runs if reinforcement were contingent upon such differential responding, and the few relevant experiments (e g , Rubinstein, 1961, Schoonard & Restle, 1961) confirm that they do To what extent dependencies of behavior upon cues from event runs develop under the conditions of the standard probability learning situation is quite another question—and one on which the literature contains little evidence

The factual matters mentioned above by no means exhaust the supply However, they are the principal items that are more or less peculiar to the standard probability learning situation Empirical relationships which are studied in much the same way in connection with probability learning as with other types of learning will be considered in later sections

Systematic Interpretations

We turn now to the question closest to the main theme of the present volume What general classes of psychological processes are represented in the experiments on simple predictive behavior? Two principal answers have been proposed, or assumed, by different groups of investigators (*a*) A relatively elementary form of learning akin to classical conditioning, (*b*) a more complex form of learning or problem solving, not analyzable in terms of simple stimulus response concepts These are not mutually exclusive possibilities, but in practice they have been generally treated as though they were It will be convenient to bring them up *seriatim* for discussion

theory of probability learning can generate accurate predictions of response probability on a given trial of a series, given appropriate information about the immediately preceding sequence of events Results in this category which have attracted the most widespread interest are those having to do with "recency curves"

Extant versions of reinforcement theory, as they have been applied to the two-choice, probability learning experiment, predict positive recency curves That is, at any point in a series of trials, the greater the number of consecutive, immediately preceding occurrences of a given event, the more likely the *S* should be to predict that event on the next trial Evidence bearing on this prediction was first reported by Jarvik (1951), and proved sharply negative Rather than following the predicted function, *S*s tended to respond in terms of the "gambler's fallacy," that is, to be more likely to predict an event the longer it had been since the last preceding occurrence of the event This "negative recency" effect was also found by a number of subsequent investigators (e g, Anderson, 1960, Feldman, 1959, Nicks, 1959), the characteristic recency curve showing an increase in probability of a response after one or two reinforcements, then a decrease as the run of consecutive reinforcements continues

A satisfactory interpretation of this phenomenon can hardly be expected until more information is available concerning its course of development and stability The suggestion has been put forward (Estes, 1962a) that the negative recency function results largely from response tendencies the *S*s bring with them to the experiment, via generalization from other situations, but which extinguish with experience in the experimental situation Some support for this interpretation has been provided by recent studies which have utilized sufficiently long series to provide a reasonable opportunity for the posited extinction to occur (Anderson, 1960, Edwards, 1961, Friedman et al, 1963) If previous analyses of the probability learning situation are extended, as in Restle's (1961, Ch 6) treatment of choice behavior, to allow for the *S*s' responding to runs of reinforcing events of different lengths as differential cues, a reinforcement theory could accommodate a transitory negative recency effect A bowed recency curve would be expected to develop early in a series simply because learning would proceed more rapidly with respect to cues from the shorter, and therefore more frequently occurring, event runs Persistence of a negative recency effect in asymptotic data would be compatible with Restle's model but very difficult to account for in terms of reinforcement principles [2]

[2] Provided that the experimental procedure avoids artifacts such as those that arise from imposing constraints on the relative frequencies of reinforcing events within short trial blocks which would lead to differential reinforcement of negative recency behavior

Another active line of investigation has been concerned with attempts to show that the effective stimulus for the *S* at any point in a probability learning series is the immediately preceding run of similar reinforcing events There is ample evidence that *Ss* respond differentially to the stimulus aftereffects of the immediately preceding trial, this is seen most directly in experiments which arrange different probabilities for a given reinforcing event depending on which event occurred on the preceding trial (Hake & Hyman, 1953, Engler, 1958, Anderson, 1960, Shelly, 1960a, 1960b) Attempts to show that *Ss* discriminate the differential stimulation associated with different lengths of runs of a given event have not yielded entirely conclusive results, partly because the investigators have not used available analytic techniques to full advantage for the purpose of partialling out the reinforcing effects of a given event run from its effects, if any, as a discriminative stimulus (see, e g , Feldman, 1959, Goodnow, Rubinstein, & Lubin, 1960, Nicks, 1959) One would surely expect that adult human subjects could, up to some limit, respond differentially to stimulation from different lengths of runs if reinforcement were contingent upon such differential responding, and the few relevant experiments (e g , Rubinstein, 1961, Schoonard & Restle, 1961) confirm that they do To what extent dependencies of behavior upon cues from event runs develop under the conditions of the standard probability learning situation is quite another question—and one on which the literature contains little evidence

The factual matters mentioned above by no means exhaust the supply However, they are the principal items that are more or less peculiar to the standard probability learning situation Empirical relationships which are studied in much the same way in connection with probability learning as with other types of learning will be considered in later sections

Systematic Interpretations

We turn now to the question closest to the main theme of the present volume What general classes of psychological processes are represented in the experiments on simple predictive behavior? Two principal answers have been proposed, or assumed, by different groups of investigators (*a*) A relatively elementary form of learning, akin to classical conditioning, (*b*) a more complex form of learning, or problem solving not analyzable in terms of simple stimulus response concepts These are not mutually exclusive possibilities, but in practice they have been generally treated as though they were It will be convenient to bring them up *seriatim* for discussion

Stimulus Response Reinforcement

The experimental situation in which most of the data on human probability learning have been obtained lends itself quite naturally to an interpretation in terms of such terms as "reinforcement," "acquisition," and "extinction", consequently it is not surprising that these terms have been more or less routinely applied to phenomena of verbal conditioning This practice has been opposed on doctrinaire grounds by investigators who find reinforcement concepts uncongenial, and on occasion investigators who habitually interpret learning experiments in these terms have raised questions as to whether the same processes are involved in verbal conditioning as in, say eyelid conditioning, or other fully accredited varieties of conditioning or elementary associative learning (e g, Hake, Grant, & Hornseth, 1951) But there seems to have been no systematic consideration of the applicability of reinforcement and related concepts

Where may we hope to find evidence on this point? An obvious possibility is to examine similarities and differences in the way experimental variables influence behavior in the probability learning situation as compared to other learning situations This approach seems likely to prove instructive, but any conclusions suggested by a review of empirical relationships should be tempered by the realization that identity of underlying processes or mechanisms need not always be closely reflected in phenotypic similarities One may put identical engines in a racing car and an airplane, but this will not guarantee similar behavior on the part of the two vehicles

Effects of reinforcing events on response probability —A prerequisite to the analysis of any situation in terms of reinforcement theory is the identification of experimental operations having the property that if one of these operations is applied on any trial, it produces an increase in probability of the reinforced response upon the next recurrence of the same stimulus situation On the negative side, it must be conceded that no reinforcing operations have been shown to satisfy this requirement without exception in the probability learning situation There are some relatively rare occasions, especially on the earliest trials of a series run with experimentally naive *Ss*, when administration of a reinforcing event actually results in a decrease in probability of the presumably reinforced response The positive evidence is much more striking, however Even in experiments which yield the negative recency effect, it is generally found that probability of a response is higher after a reinforced than after a nonreinforced trial (Jarvik, 1951, Nicks, 1959, Anderson 1960) When response proportions are plotted for a group of *Ss* all of whom receive the same sequence of reinforcing events over a number of successive trials, the ups and downs of the response curve follow the event sequence so faithfully that this phenomenon has suggested

the designations probability tracking (Estes 1957b) or probability following (Edwards 1961)

The nature of reinforcing events—In Humphreys arrangement of the verbal conditioning experiment occurrence or nonoccurrence of the event predicted by the *S* was supposed to correspond to the occurrence or non occurrence of the unconditioned stimulus in a classical conditioning experi ment How deep the correspondence goes is problematical Little is known unfortunately concerning the *S*'s original response to the reinforcing signal in the verbal conditioning situation Some investigators (Estes & Straughan 1954) have assumed that the reinforcing signal evokes a tendency upon the part of the *S* to make possibly covertly the response indicated to be correct If the *S* has predicted correctly on a given trial the signal would lead him to repeat the same response whereas if he has predicted incorrectly the signal would lead him to make a corrective response Attempts to get direct evidence relative to these response tendencies by recording muscle action potentials (Kent 1958) have yielded results that may perhaps be considered mildly positive but the MAP patterns are difficult to interpret in any clear cut fashion There is no reason to believe that any such response tendencies are unlearned but this is not a critical point for the unconditioned re sponse in a classical conditioning experiment may be one established by previous conditioning The *S*'s response to the reinforcing signal may certainly be modified by instructions but again the same is true of UCRs in at least some cases Apparently there is little prospect of obtaining cogent independent evidence to support the assumption of covert corrective re sponses There remains the possibility that continued use of this assumption in theories of probability learning may generate convincing indirect evi dence one way or the other depending on the success of the theories

Comparing the standard probability learning experiment with other human learning experiments we find general similarity with respect to the reinforcing operations used and also with respect to some important equiva lences In paired associate experiments one obtains about the same results when reinforcing operations supply information only (by paired presenta tions of the stimulus and correct response members of each item) as when they involve positive and negative aftereffects (confirmation of correct anticipations and correction of incorrect ones) Thinking in terms of law-of effect notions one might expect that generally in the probability learning situation the reinforcing effect of an event upon the response of predicting it would be different depending on whether or not the response had oc curred on the given trial That is denoting the two reinforcing events by A_1 and A E_1 and E and the responses of predicting these events by A_1 and A respectively the event E_1 will sometimes occur in the sequence A_1—E_1 and sometimes in the sequence A—E_1 In both cases the information given the

S would be the same, but in the first instance one might consider the A_1 response to have been rewarded, whereas in the second instance this response could not have been rewarded, in the usual sense at least, since it did not occur If the effect of the E_1 event were greater in the first instance, we would have support for an interpretation in terms of the law of effect, or perhaps confirmation of expectancies If the effect of the E_1 event were the same in the two instances, we would have support for the assumption that reinforcing events in probability learning function similarly to unconditioned stimuli in classical conditioning

It is not easy to find evidence bearing relatively directly on this issue Firstly, not many investigations have reported data showing detailed dependencies of responses on preceding responses and events Secondly, even given suitable data, there are as yet unsolved problems of interpretation To illustrate one approach, let us examine results reported by Suppes and Atkinson (1960, Ch 10, Group Z) for a group of *S*s run in a standard, two-choice predictive experiment. In order to determine the effect of the reinforcing event in each type of trial, we must control for the preceding response Thus we are interested in the differences

$$P(A_1\ E_1A_1) - P(A_1\ A_1)$$

and

$$P(A_1\ E_1A_2) - P(A_1\ A_2),$$

where the first line is to be read " the proportion of occurrences of response A_1 following the sequence A_1—E_1 on the preceding trial minus the proportion of occurrences of A_1 following an A_1 on the preceding trial (regardless of the reinforcing event)," and the second line similarly Values reported by Suppes and Atkinson for the last 100 trials of a 240-trial series yield differences of 07 in both cases A similar analysis is available for a group of 80 *S*s run through a series of 48 trial blocks with varying probabilities of reinforcement (Friedman et al , 1963) Using the pooled data for the last 11 trials of each of 10 blocks run with a .50 50 schedule (these blocks were interspersed between blocks run with other schedules), the two differences prove to be

$$P(A_1\ E_1A_1) - P(A_1\ A_1) = 616 - 523 = 09$$

and

$$P(A_1\ E_1A_2) - P(A_1\ A_2) = 532 - 432 = 10$$

These results accord nicely with expectations based on the assumption that a given reinforcing event produces the same effect regardless of the response it follows on a particular trial Unhappily, we have no assurance that they could not also follow from various combinations of assumptions specifying different amounts of learning on "correct" and "incorrect" trials So far as

I have been able to determine, no model free method of analyzing the data from simple, two-choice, noncontingent schedules yields unequivocally different predictions for the two types of assumptions There remains the possibility of testing alternative hypotheses about the action of reinforcing events within the framework of specific learning models This strategy will be illustrated in a later section

Rewards such as monetary payoffs appear to function in probability learning situations much as in other human learning situations, either verbal or perceptual motor, when administered contingent upon correct responses (see, e g , Goodnow, 1955, Robillard, cited in Bush & Mosteller, 1955, Siegel & Goldstein, 1959, Suppes & Atkinson, 1960), and the same appears to be true of reinforcement by escape or avoidance of unpleasant or otherwise undesirable stimulation (Brody, 1957, Detambel & Stolurow 1957, Straughan, 1956) Little is known about the role of parameters of reward, however Delay of reward seems not to have been studied at all in the probability learning situation Magnitude of reward appears to influence primarily the asymptote of performance rather than the rate of approach to asymptote (Siegel & Goldstein, 1959, Suppes & Atkinson, 1960, Ch 10, Bush & Mosteller, 1955, Ch 13, Taub & Myers, 1961), a relationship which has been frequently observed in studies of simple instrumental or trial-and error learning and which dictated Hull s (1943) postulate concerning magnitude of reward

There is no room for doubt that in any of the standard human learning situations administration of a suitable reward following occurrences of a designated response is a sufficient condition for certain changes in performance which are taken to be symptomatic of learning I would suggest, however, that insufficient critical examination has been given the question as to whether the usual temporal relation between response and reward is a *necessary* condition for these changes, as assumed in the older Thorndikian, as well as the contemporary drive reduction, interpretations of reward

In addition to satisfying a drive or motive, if it does, the occurrence of a reward also conveys information Which aspect is responsible for the reinforcing function of the reward? In the few studies addressed specifically to this problem, it has proved surprisingly difficult to demonstrate any effect of a reward as a "satisfying aftereffect" beyond that it exerts as an informative signal (Bitterman, 1956, Hillix & Marx, 1960) It is interesting to note in this connection that variation in ambiguity of an informative reinforcing signal has effects very similar to those of variation in magnitude of reward (Estes & Johns, 1958) Analyzed in stimulus response terms, the occurrence of a reward involves a stimulus whose response may become conditioned to cues preceding it and thus become anticipatory, for example, an approach or avoidance response evoked by a particular reward could

become conditioned to cues associated with a response manipulandum If the function of a reward were basically the same as that of an informative signal or an unconditioned stimulus, it would not be essential that a reward be administered immediately following the response undergoing learning One way in which the usual contingency can be broken is illustrated by a study recently completed in the Indiana Laboratory (M Cole, L Keller, C J Burke, & W K Estes, unpublished) Each trial began with presentation of some one member of a set of discriminative stimuli, to which the *S* responded by choosing one or the other of two response buttons Then numerals appeared in each of two windows, showing the payoff associated with each response button on that trial, thus the *S* had full information on each trial, although he received a given payoff only if he had operated the corresponding response button It is hoped that suitable experimental analyses in this situation may provide evidence on the critical question as to whether the effect of presenting a given reward stimulus is dependent on or independent of its temporal relation to the rewarded response

Effects of nonreinforced trials—By "nonreinforced," or "blank" trials in the probability learning situation we mean trials on which none of the experimenter-defined reinforcing events occurs Owing to a fortuitous property of the experimental situation it was not possible in the earliest two-choice probability learning studies (Humphreys, 1939, Grant, Hake, & Hornseth 1951) to distinguish operationally between nonreinforcement of one response and reinforcement of the alternative The *S*'s task in those studies was to predict whether a reinforcing light would or would not occur on each trial During "acquisition" the light was scheduled to appear on, say 100% or 50% of the trials, depending on the condition for a particular group, then during "extinction" the light no longer appeared at all It was said that the response of predicting the light received 100% or 50% reinforcement during acquisition and no reinforcement during extinction However, failing to predict occurrence of the light was equivalent to predicting its nonoccurrence, and the symmetry of the situation was such that one would expect nonoccurrence of the light to have the same reinforcing effect on the latter response that occurrence of the light had on the former, extinction of one response could not be distinguished from acquisition of the other

In the experimental variation introduced by Estes and Straughan (1954), the symmetry of the responses and reinforcing events was made explicit in a way which has become standard for most subsequent work A separate reinforcing light was associated with each response key, and the *S*'s task was to predict which of the two lights would occur on each trial In the modified situation, extinction and reversal can be distinguished, for following a series in which a given response has been the more frequently reinforced, one may

shift either to a reversal series in which the light associated with the alternative response occurs more frequently or to an extinction series in which neither of the reinforcing lights occurs

To my knowledge, the first use of nonreinforced trials in the symmetric choice situation occurred by accident in an experiment from which the acquisition data have been reported by Burke, Estes, and Hellyer (1954) It had been intended to follow the acquisition series by a reversal, as in a later study by Kalman (1961), but through a failure of communication, the *Ss* were instead shifted onto a series of nonreinforced trials To the considerable surprise of the investigators, the mean response probabilities, approximately 90 10 at the end of acquisition, remained virtually constant over the extinction series Essentially the same result was obtained in a later study by Neimark (1953) when a block of 50 nonreinforced trials was given following a 66 34 noncontingent acquisition series These observations suggested that under some circumstances blank trials might be neutral events, i e, produce no change in response probabilities This hypothesis received clearcut support from a study by Atkinson (1956) in which blank trials occurred with various probabilities following each response alternative throughout acquisition

On the assumption that the reinforcing effects of various trial outcomes depend on the responses evoked from the subject when particular events occur (or fail to occur), we might expect the function of blank trials to be quite sensitive to the experimental context and instructions Under experimental conditions differing so slightly from those of the studies mentioned above that the critical differences have not yet been isolated other investigators (Anderson & Grant, 1957, Greeno, 1962) have found blank trials to produce an effect, probably small, relative to the effects of reinforcing stimuli, which may be described as 'decremental' in the sense that it involves a regression of response probabilities toward a chance level The whole picture is reminiscent of that in the area of eyelid conditioning where blank (unconditioned stimulus omitted) trials occurring during acquisition usually exert a decremental effect on conditioned response probability although in some cases this effect is very small (Grant & Schipper, 1952)

When the experimental context is such as to lead the *S* to interpret the absence of a reinforcing signal following one response to mean that the alternative response would have been correct, particularly likely to be the case in contingent situations (Brand, Woods, & Sakoda, 1956, Brand Sakoda & Woods, 1957, Neimark, 1956), and even more reliably so when instructions are designed to evoke this interpretation (Koehler, 1961) the omission of a reinforcing signal appears to produce an effect equal and opposite to the effect of that signal Under these circumstances, a shift to extinction, in the sense of a block of successive blank trials, following an

acquisition series results in response probabilities going to a chance level at a rate about equal to that of original acquisition (Neimark, 1953) Although there remains adequate scope for more detailed analyses, prospects appear reasonably good that the whole problem of effects of nonreinforced trials can be given a coherent interpretation in terms of the responses evoked by reinforcing stimuli and their omission under various experimental arrangements

Intertrial interval—Evidence concerning the role of trial spacing in human probability learning is rather sparse Studies in which intertrial intervals have been varied over about the same range usually studied in conditioning experiments have generally reported no significant effects on rates of acquisition or extinction (reversal) (e g , Anderson, 1960, Grant Hornseth & Hake, 1950) On the assumption that this insensitivity might result from *Ss'* bridging the temporal gaps between trials by verbal rehearsal or persisting postural adjustments, other investigators (Estes, 1955, Straughan 1956) have used distracting tasks to fill the intervals, under these conditions, rate of learning and rate of reversal have both proved to be inversely related to intertrial interval over a range of 5–30 sec No evidence has been forthcoming to indicate that massing of trials in the two-choice situation produces a decrement in the terminal level of performance comparable to that routinely observed in classical conditioning experiments One must note, however, that even if the same process were operating its effects might not be manifest in choice data since presumably both of the alternative responses would be affected Relevant information might be provided by studies in which intertrial interval was varied but with latency as well as choice data recorded In the only published study of probability learning involving well spaced blocks of trials run under the same reinforcement schedule, a spontaneous regression effect, comparable to that familiar in simple instrumental and classical conditioning situations, was observed (Estes Burke, Atkinson, & Frankmann 1957)

Resistance to extinction—Two of the classical problems in both Pavlovian and instrumental conditioning have been those of the effects of number of reinforcements and intermittency of reinforcement upon resistance to extinction The first of these relationships has received little attention in studies of two-choice behavior The only relevant experiment I have come across is one reported by LaBerge (1959a) in which different groups of *Ss* were given different numbers of trials on a 50 50 schedule before being shifted to a 10 90 schedule The rate at which the learning curves went from a mean probability of 50 to the new asymptote of 10 for the less frequently reinforced response was inversely related to the number of pre shift trials on the 50 50 schedule quite in accord with the usual finding in conditioning studies

The effect of partial vs continuous reinforcement was one of the first problems studied in the verbal conditioning situation (Humphreys, 1939, Grant, Hake, & Hornseth, 1951), and the gross results, increased resistance to extinction with partial reinforcement and a direct relationship between percentage of reinforcement of a response over a considerable range of values and speed of extinction or reversal, have, again, been in agreement with the relationships generally found to obtain for other conditioning situations At a finer level of analysis, Hake, Grant, and Hornseth (1951), studying verbal conditioning and eyelid conditioning with parallel experimental designs, found similar relationships between resistance to extinction and the number of shifts from nonreinforcement to reinforcement during acquisition They found a difference in the effect of length of nonreinforced runs during acquisition in the two situations, but with some question as to the significance of the difference

Unusually great resistance to extinction bas been reported by Rogers, Webb, and Gallagher (1959) for groups run in a two choice situation with partial knowledge of results, under this condition, *S*s assigned the task of predicting whether or not a light would appear received the light stimulus with a given probability (30, 50, and 85 for different groups) on trials when they predicted the light but received no information on trials when they did not The authors point out the analogy between this result and the familiar extreme prolongation of extinction following partial reinforcement in operant conditioning situations, where the animal, if it does not press the bar or peck the key, receives no information as to whether the response would have been followed by reinforcement

Quantitative analyses—Although the comparisons summarized above are not unequivocal on all points, it appears that on the whole the effects of most experimental variables on probability learning are similar to their effects on other varieties of human learning The correspondences are close enough in a substantial number of respects to lend some encouragement to the assumption of common basic processes Unfortunately, however, differences in the behavioral measures used in different situations, together with differences in the operations used to produce presumably similar effects, limit the possibilities of arriving at any firm conclusions from comparisons of gross empirical relationships alone An auxiliary source of relevant evidence which may he worth considering would involve comparisons at a more analytical level, as permitted by the development of descriptive and predictive theories in the various areas Once we have established satisfactory theoretical models for each of the principal 'types' of human learning, comparisons of these may reveal communalities that would never be apparent at a strictly phenotypic level As a first step toward a comparative analysis, I shall now turn to a brief resume of a theory which appears

to offer some promise for this purpose in the area of probability learning.

The theory to be discussed is a particular case of stimulus sampling theory which may, for brevity, be designated as the *pattern model* (Estes, 1957a, 1959b). I am utilizing this model, not because it is even a close approximation to the most adequate that could be offered as a situational theory for probability learning, but because it appears to offer the best available compromise between the always conflicting demands of simplicity and predictive power. The principal assumptions are as follows.

Response analysis. The theoretical dependent variables are the probabilities of the response alternatives, which may be denoted A_1, A_2, etc., recognized by the experimenter. It should be noted, however, that although these probabilities are evaluated by the observed relative frequencies of key presses, yeses and noes, or the like, there is no assumption that a particular response alternative, A_i, is to be empirically identified solely with a particular movement such as pressing a given key or saying "yes." It is possible, for example, that A_i might represent a class of verbal hypotheses, or rules, all of which have in common the property of specifying a given verbal response or operation of a given response key in a particular stimulus situation.

Stimulus analysis. Recognizing that, even though the experimenter presents the same signal at the beginning of each trial, the *S* may discriminate stimulus traces from the outcomes of one or more preceding trials, it is assumed that in general the *S* may have potentially available for sampling a population of different stimulus patterns. Both the number, N, and the nature of these stimulus patterns will be determined jointly by the stimulus displays presented by the experimenter at the onsets of successive trials, by the reinforcement schedule, and by the learning history of the *S* prior to the given experiment. It is to be expected that a fully adequate account of a particular experiment would require identifying all of the stimulus patterns actually discriminated by the *S* and determining their sampling probabilities. But for our present purposes it will be expedient to limit attention to the simplest special case, in which all stimulus patterns are assumed to have equal sampling probabilities and no attempt is made to identify them individually.

Learning assumptions. In the simple pattern model, it is assumed that at any time each stimulus pattern is associated with ("conditioned to," or "connected to") exactly one of the response alternatives A_i, and that when a pattern is sampled, the response associated with it is necessarily evoked.[3]

[3] In a more general case of the theory than that considered here, allowance would be made for the possibility that, at the beginning of learning, some or all of the stimulus patterns might be "neutral," that is conditioned to none of the alternatives A_i (LaBerge, 1959b). On trials on which neutral patterns were sampled, the *S*

Thus the probability of a given response at any time is equal to the proportion of all available patterns that are associated with it Learning is assumed to occur on an all-or none basis Specifically, if A_i is the response of predicting the reinforcing event E_i then an occurrence of E_i is assumed to evoke from the *S* a response, possibly covert, of class A_i, and with some fixed probability *c* this response becomes conditioned to the stimulus pattern sampled on the trial (if it is not already so conditioned) It will be noted that the learning process assumed is at the opposite extreme from that implied by the traditional law of effect or by a drive reduction interpretation of reinforcement, for learning occurs only on 'incorrect' trials that is trials of the type A_1E_2 or A_2E_1—no learning occurs on trials when the *S* s prediction is confirmed (A_1E_1 or A_2E_2)

Before proceeding toward data, it might be well to insert a few remarks concerning the relationship between the pattern model and the familiar linear model [4] which has been applied to numerous studies of probability learning, beginning with the analysis by Estes and Straughan (1954) In the original, and simplest, form of this model, it is assumed that on each trial when a given reinforcing event E_i occurs, the probability of the corresponding response, A_i, receives the increment described by the linear transform

$$p_{n+1} = (1-\theta)p_n + \theta, \tag{1}$$

where p_n and p_{n+1} denote the probabilities of response A_i on trials *n* and $n+1$, respectively, and θ is a learning parameter with a value between 0 and 1 The effect of the reinforcing event depends on the current level of response probability, but it is entirely independent of the response which happens to occur on the given trial

Important similarities and differences between the pattern and linear models arise from the fact that in the latter, Eq 1 describes the change in response probability on a reinforced trial for an individual *S* and also the mean change for a population of *S*s, whereas in the former, learning in individual *S*s occurs by jumps in *p* values from 0 to 1 (or 1 to 0) and Eq 1 applies only to the mean change for a population of *S*s [for full elucidations

would choose at random from the set of response alternatives and in this event learning could occur on correct trials Over a series of trials, all of the available patterns would be conditioned to members of the response set (A_i) and then the model would reduce to the special case considered in this paper It might be mentioned that the treatment of response latencies in the probability learning situation within the stimulus sampling theory involves the notion of neutral elements (La Berge 1959b)

[4] For variations and extensions of the linear model see Anderson 1959 Bryant and Marica (1959) Burke and Estes (1957) Bush Mosteller and Thompson (1954) Hanania (1960) Overall (1960) Radlow and Segel (1960) Restle (1957) Siegel (1961) and Sternberg (1959a 1959b)

of both models, see Estes (1959b) and Estes & Suppes (1959)] Because of the similarity in the one respect, both models yield identical predictions concerning mean learning curves and asymptotes and curves of response probability as a function of runs of reinforcing events (recency curves) in the noncontingent case, and both yield general probability matching theorems (Estes, 1957a, 1957b) Because of the difference in the other respect, the models yield different predictions concerning sequential dependencies of responses upon combinations of preceding responses and reinforcing events However, as we shall see shortly, even where there are differences in predictions at the level of fine structure of the data, these differences are apt to be very small When one is concerned specifically with evaluating all or none vs incremental conceptions of the basic association process, the differences, though small, may be of considerable importance But for all practical purposes they may be ignored while we proceed with the task immediately at hand to determine how well probability learning data support a theory (or, more precisely, class of theories) based on the concepts of stimulus, response, and conditioning by contiguity

When first applied to the two-choice probability learning situation, statistical learning models seemed quite successful by the standards prevailing at the time They provided satisfactory descriptions of the typical form of the learning curve, not a taxing assignment, but still a test that had to be passed by any theoretical candidate The parameter θ was not invariant over all combinations of initial response probabilities and reinforcement probabilities as required by the simplest model, but continued research yielded normal progress toward clearing up these disparities The predicted sensitivity of response probabilities to local fluctuations in density of reinforcing events was nicely borne out (Estes & Straughan, 1954) And, most striking to the general public, the predicted matching relationship between asymptotic response probability and probability of reinforcement proved able to account for something of the order of 98% or 99% of the variance in terminal mean response proportions for a large number of groups run under various noncontingent schedules in different laboratories (Estes 1961b, 1962a)

The cogency of these lines of evidence in apparent support of the statistical theory might be questioned on various grounds In connection with the asymptotic predictions one cannot overlook the fact that deviations from predicted values are sometimes observed (evidence on this point has been reviewed in an earlier section) and an obvious corollary, that in these instances other models can be constructed which do better (Edwards 1956, Gardner, 1957) Conversely it may be contended that even if probability matching does occur it can be predicted (for the case of simple, noncontingent reinforcement schedules) from other theories (Feldman &

Newell, 1961, Simon, 1956) The statistical theory appears to be nicely caught in a logical crossfire

The one question I would raise in connection with these observations is that of whether they necessarily constitute negative evidence relative to the theory The stimulus sampling model was not developed as a situational theory for any particular type of probability learning experiment, and its concepts and assumptions do not include explicit reference to all of the variables known to influence probability learning Consequently, it could hardly fail to be the case that for almost any particular experiment some specially constructed model would yield more accurate predictions Nevertheless, if the statistical theory correctly represents the mode of operation of certain variables common to all probability learning situations, then this theory should be able to accomplish something that the various local theories cannot—namely, account for a major portion of the variance of the data over a wide range of experimental arrangements Let us summarize the results of a few of the instances in which the statistical model has been applied to new experimental situations (that is, ones involving conditions that had not been investigated at the time of original formulation of the model)

The first experimental variation to appear involved simply making the probabilities of reinforcing events depend on the *S* s responses (e g, Estes, 1954, Neimark, 1956) Whereas in the noncontingent case the probability of a given reinforcing signal is the same on all trials, in the contingent case the probability of a particular signal might be, say, 85 following an A_1 response but 30 following an A_2 Now it is hard to see how the *S*s, unless they were rather more skilled at mathematics than most of our undergraduates and also were allowed ample time for pencil and paper computations, could manage by any rational scheme to match their response probabilities to the relative frequencies of reinforcement, for now the final level of probability for a given reinforcing event depends jointly on the contingencies set up by the experimenter and the sequence of responses produced by the *S*s Nonetheless, the stimulus sampling theory predicts that on the average the *S* will adjust to the variations in frequencies of the reinforcing events resulting from fluctuations in his response probabilities in such a way that his probability of making a given response will tend to stabilize at the unique level which permits matching of the response probability to the long term relative frequency of the corresponding reinforcing event Further, this matching value can be predicted a priori from the theory, given only the experimenter's reinforcement contingencies

Interpretation of results obtained with the contingent procedure is complicated by the fact that in nearly all cases the change from noncontingent to contingent probabilities has been accompanied by other changes in the

experimental context or instructions. I think that perhaps the reason for this confounding is that investigators have tended to think of the contingent case as analogous to instrumental conditioning, and of the noncontingent case as analogous to classical conditioning. I doubt that the analogy goes very deep, but a consequence of this orientation is that in contingent experiments *Ss* have generally been instructed, not simply to indicate their expectation as to which event would occur on each trial, but to try to produce a particular event, e.g., a red light (Brand, Woods, & Sakoda, 1956), which then serves as a reward. A difficulty with this procedure from the standpoint of a stimulus response contiguity analysis is that we do not know how the *S* will respond to omission of the reinforcing light. Application of the statistical model is predicated on the assumption that a corrective response will be evoked when the light does not appear, that is, if the *S* makes an A_1 response and the light does not come on, he will tend, overtly or covertly, to make an A_2. But in nearly all experiments with the contingent procedure, *Ss* have not been instructed that nonreward of one response means that the other would have been correct, and in these cases there has generally been a tendency for terminal proportion of the more frequently rewarded response to go somewhat above the matching value (Detambel, 1955, Koehler, 1961, Woods, 1959a, 1959b). When instructions have been designed to lead to corrective responses on nonrewarded trials, probability matching has been approximated quite closely, at least for a couple of hundred trials (Koehler, 1961).

Even in experiments where numerical predictions are not as close to the mark as might be desired, it is still reasonable to ask whether the theory accounts for the main effects of systematic variations in reinforcement contingencies. To illustrate, let us consider an experiment conducted by Brand, Woods, and Sakoda (1956) to investigate the effects of varying differences, as opposed to ratios, between the probabilities of reinforcement in a two-choice, contingent situation. With probabilities of reward for the more and less frequently rewarded response denoted by π_1 and π_2, respectively, the rows in Table 3 correspond to the differences, $\pi_1 - \pi_2$, and the columns correspond to the ratios π_1/π_2, which were combined in a factorial design.

This problem exemplifies a type which poses a nice test of the usefulness of a theory, for one could hardly expect to assess the relative importance of the two factors accurately on intuitive grounds. According to their account, the authors largely on the basis of Brunswik's (1939) analysis of probability learning, expected in advance of the experiment that the *Ss'* behavior would be about equally sensitive to variation in ratios of reward probabilities and to variation in differences with ratios held constant. Specifically, probability of the more frequently rewarded response should increase as a

TABLE 3

OBSERVED AND THEORETICAL PROPORTIONS OF MORE FREQUENTLY REWARDED RESPONSE IN RELATION TO RATIOS AND DIFFERENCES BETWEEN REWARD PROBABILITIES

		Ratio			
	Difference	2 1	3 1	6 1	11 1
Observed[a]	50	92	81	74	75
	40	80	78	73	76
	30	60	55	64	61
	20	62	63	60	57
Theoretical	50	1 00	75	69	68
	40	75	67	64	63
	30	64	61	59	59
	20	57	56	56	56

[a] Read from graphs in Brand Woods, & Sakoda (1956)

function of both increasing differences and increasing ratios Also, on the hypothesis of a probability discrimination threshold, a significant interaction was expected between these variables The results of the experiment were quite opposed to these expectations Response proportions for the final 40 trials of a 120-trial series (see Table 3) show highly significant covariation with differences in reward probability at all ratios, but virtual constancy over a wide range of ratios, especially at the smaller differences, and little indication of an interaction Where a change in response probability did occur as a function of reward-ratio, the change was in the opposite direction to that predicted from Brunswik's analysis Why should this be? According to the statistical theory, it is simply the result to be expected if learning is proceeding in accord with the assumptions discussed earlier Computing theoretical asymptotes for the combinations of reward probabilities used in this study we find that the model generates the, to my mind highly counter-intuitive, prediction of a *decrease* in probability of the more frequently rewarded response as the ratio of its reward probability to that of the alternative response *increases*, most of the drop occurring between ratios of 2 1 and 3 1, with virtual constancy from the latter value out to 11 1 (the largest ratio tested) The result observed was a decline in the marginal means from 74 to 69 between 2 1 and 3 1, the corresponding a priori predictions being 74 and 66, then a drop of only 02 between 2 1 and 11 1, to be compared to a predicted drop of 04 However, as the difference in reward probabilities increased from 20 to 50, the mean proportion of occurrences of the more frequently rewarded response increased in

a positively accelerated fashion from 60 to 80, which may be compared with the prediction of a positively accelerated increase from 57 to 79

This example has seemed worth going through in some detail because it helps make the point that the most instructive tests of a quantitative theory have to do with its ability to account for intrinsically interesting functional relationships rather than simply the predicting of absolute values of data statistics for their own sakes Further, re examining the experimental arrangement in the light of the theory, we can gain some insight into the reason for the rather unexpected result According to the theory, learning occurs primarily on trials when incorrect responses occur Consequently, the key question is how the frequency of incorrect occurrences of each response will be affected by variation in differences or ratios of reward probabilities With respect to differences the situation is relatively straightforward Increasing the difference reduces the probability of incorrect occurrences of the more frequently rewarded response, or increases this probability for the other response, or both simultaneously—in any event leading to an increased difference between the two response probabilities Effects of increasing the ratio of reward probabilities may be more complex Suppose we start with a schedule, say 1 0 50, under which the favored response has a high asymptotic probability (in this case 1 0) and then proceed to increase the ratio of reward probabilities from 2 1 to 3 1 while holding the difference constant, by shifting to a 75 25 schedule We will have changed the probabilities of nonrewards following each response by equal amounts, but the response which had higher probability under the first schedule will on that account be subject to the greater increase in frequency of nonrewards—and therefore the greater reduction in response probability Further increases in the ratio of reward probabilities entail progressively smaller changes in frequency of nonrewards (going from 3 1 to 6 1 increases the probability of nonreward following each response by less than half as much as going from 2 1 to 3 1) and consequently progressively smaller changes in probabilities of the two responses

Similar analyses of more complex reinforcement contingencies have been published elsewhere, and I shall not attempt to review all of them Among these are the problem of contingency with lag (Estes, 1957b, 1959a), combinations of escape and avoidance contingencies (Brody, 1957), and contingency of one S's reinforcement upon the behavior of another (Atkinson & Suppes, 1959, Burke, 1959, 1960, Estes, 1957a, Suppes & Atkinson, 1960) I wish only to emphasize that each case involves not only the testing of numerical predictions but also the elucidation of more general relationships One can, for example, account for the curious finding (which no one had thought to look for until it was predicted by the theory) that if the probability of a reinforcing event on any trial depends on the response made by the subject ν trials earlier, then, other things equal, the asymptotic

response probability is independent of the lag v. Similarly, one can show in detail how it is that some schedules of outcome contingencies ("payoff matrices") in simple games of strategy permit the *S*s to approximate game-theoretic optimal strategies while other schedules lead to large deviations from the optima.

Any attempt to appraise the stimulus sampling interpretation (or, for that matter, any other interpretation) of probability learning must consider also the ability of the theory to predict details of the finer structure of the data. These details include, e.g., dependencies of responses upon particular sequences of responses and reinforcing events over the immediately preceding portion of the trial series. As might be expected by analogy with other research areas, published results pertaining to this type of prediction exhibit a by no means random pattern of failures and successes. Over the collection of pertinent studies (including Anderson, 1960, Brody, 1958, Estes & Straughan, 1954, Friedman et al., 1963, Nicks, 1959, Suppes & Atkinson, 1960), the proportion of successes has been closely correlated with the number of measures taken in individual experiments to satisfy the simplifying assumptions of the particular case of the theory used for prediction.

What is one to make of the fact that disparities between theory and data at the level of fine structure sometimes occur in the same studies which yield confirmatory results relative to grosser trends such as curve forms or mean asymptotes? There are dangers in hastening to dismiss the gross correspondences as "fortuitous" (cf. Anderson, 1960), for one cannot realistically expect any theory to predict all details of relevant data. Rather, it seems wise to defer judgment in cases of this sort until further analysis has provided a satisfactory interpretation of both the successes and failures in the original predictions.

Errors of prediction are always a symptom that something is wrong. And the something wrong may be the basic assumptions of the theory under test. This possibility cannot be ruled out in the case of the theory of probability learning. Conceivably the assumptions are incorrect and the correspondences between predicted and observed terminal response levels in the noncontingent case are purely accidental, the cases of observed matching having arisen through some process quite different than that represented in the statistical model. But this argument becomes rather difficult to take seriously when asymptotic predictions receive a substantial proportion of close confirmations as we continue to conduct tests under new variations in conditions. Even if second order deviations from the theory are being averaged out, it is asking a good deal of coincidence to argue that simply by chance they should often average out to the particular values predicted.

Errors of prediction must be expected to occur also in cases when the

assumptions of a theory are essentially correct, as far as they go, but when our analysis of the empirical situation has not yet isolated all of the significant sources of variation in the data. To the extent that empirical variables represented in the theory are independent of those not represented, effects of the latter may be expected to average out when predicted relationships involving the recognized variables are being tested. When there are interactions, on the other hand, the effects of the unrecognized variables will be reflected in errors of prediction.

In raising the question as to whether a particular model for probability learning is correct, I am using the notion of "correctness" only in this somewhat special sense. On the whole, the trend of recent experimental results suggests that, at least for adult human *S*s, the probability learning situation may be too complex to be handled in all desirable detail by any manageable model. Nonetheless, it is worthwhile to locate, if we can, models which accurately represent some of the component processes. Such models may provide a basis for illuminating comparisons between probability learning and other varieties of learning which are equally refractory to complete analysis.

Before we turn to comparisons involving other learning situations, it may be instructive to give one example of sequential analysis of probability learning data under relatively optimal circumstances. For this purpose I shall discuss a case in which the theory is used first to specify suitable test conditions and then to generate quantitative predictions.

It is assumed in the stimulus sampling theory that the *S* learns by the formation of associations between his choice responses and stimulus patterns which may include stimulus traces or verbalizations based on preceding events. Thus until more is known about how the subject samples stimuli based on the preceding sequence, we cannot expect much from fine grain analyses, except in the noncontingent case, where all available stimulus patterns, no matter how constituted, must have the same correlations with subsequent reinforcing events. Also, we must expect that until learning has occurred, responses to stimulus patterns available for sampling in a given situation will be determined by generalization from other situations. The most favorable circumstances for successful predictions of sequential data from the simplified model we have considered above must evidently be associated with essentially asymptotic data obtained under a noncontingent schedule with reasonably well practiced *S*s. A sample of data which seems to meet these specifications is the second 96 trials of the .8 series of the three-session study of Friedman et al. (1963) cited earlier. In Table 4 are presented the observed proportions of occurrences of triplets of the type $A_nE_nA_{n+1}$ pooled over 80 *S*s and 96 trials, together with theoretical values computed from the linear model of statistical learning theory and

TABLE 4

OBSERVED AND THEORETICAL TRIPLET PROPORTIONS FOR TRIALS 97 192 OF 8 SERIES

Trigram *n* A	Trigram *n* E	Trigram *n*+1 A	Observed value	Linear model	Pattern model
1	1	1	567	560	567
1	2	1	116	116	118
2	1	1	103	106	104
2	2	1	071	071	070
1	1	2	081	082	082
1	2	2	043	045	044
2	1	2	054	052	053
2	2	2	019	019	019

from the pattern model [for the computational formulas, see Estes & Suppes (1959) for the linear model, and Estes (1959b) for the pattern model]

Evidently we may conclude that, under conditions carefully chosen so as to minimize extraneous factors, the dependence of the S's response upon the event and the response of the preceding trial is described exactly by the theory One might wonder how the two versions of the theory can agree so closely with each other when they differ sharply with respect to one aspect of the learning process The pattern model, it will be recalled, assumes that associations form on an all or none basis, and only on incorrect trials, whereas the linear model allows changes in response probability on all trials However, it should be noted that the learning rate in this experiment was rather high, the θ value for the linear model being estimated at 15, and in this case the processes envisaged in the two models do not work out very differently The increment in response probability prescribed for any reinforced trial by the linear model depends on the current level of probability, according to the relation $p_{n+1}-p_n = \theta(1-p_n)$, and thus will be relatively large only when p_n is small, negligible when p_n is near unity Given a relatively large θ value, the response probability for any S will fluctuate considerably from trial to trial, and on the average the choice whose probability is p will occur most often on trials when p is high Therefore, correct trials (A_1E_1, if the A_1 choice is the one with probability p) will occur most often when the momentary value of p is high and the possible increment from an effective reinforcement small, but correction trials (A_2E_1) will occur most often when the momentary value of p is low and the possible increment from an effective reinforcement large Thus, al-

though for different reasons, both forms of the theory actually agree in predicting that learning will occur largely on correction trials, not because the effectiveness of the reinforcing event depends on the preceding response, but because the momentary value of response probability is correlated with the response occurrence The results shown in Table 4 seem to bear out the common prediction

The one feature of the comparisons in Table 4 that is not entirely satisfying is that some of the observed data (3 of the 7 degrees of freedom in the 8 proportions) were used to estimate the parameters of the models, and it is hard to judge by eye how much of the agreement between computed and observed values is attributable to this constraint One way to get away from this limitation, and obtain an exceedingly stringent test of the theory, is to use the parameter values already estimated from the triplets as a basis for predicting the probabilities of longer sequences These computations have been completed for all 32 5-tuplets of the form $(A_nE_nA_{n+1}E_{n+1}A_{n+2})$, as well as selected longer sequences, using the pattern model, the linear model, and a number of other cases of the general stimulus sampling theory, as part of a detailed analysis now being prepared for publication by the writer in collaboration with Professor Patrick Suppes

Another way to gain some perspective relative to the correspondences between theory and data in Table 4 is to examine predictions for the same data generated by models having the same general structure and the same number of free parameters but differing with respect to one or more assumptions Since both models represented in Table 4 assume that learning occurs primarily on error (correction) trials, it is of especial interest to consider a model which assumes instead that learning occurs primarily on correct (rewarded) trials A convenient choice for this purpose is a special case of the general linear model which has been analyzed extensively by Bush and Mosteller (1955, Ch 8) Letting p_n denote probability of the A_1 response on trial n, as before, we can state the learning assumptions for this model as follows

If A_1 and E_1 occur on trial n,

$$p_{n+1} = (1-\theta)p_n + \theta,$$

if A_1 and E_2 or A_2 and E_1, occur on trial n,

$$p_{n+1} = p_n,$$

if A_2 and E_2 occur on trial n,

$$p_{n+1} = (1-\theta)p_n$$

Thus probability of A_1 undergoes an increment on trials when A_1 occurs and is correct, and a decrement on trials when A_2 occurs and is correct,

whenever either response is incorrect response probability remains unchanged

From the standpoint of theories which take satisfying aftereffects or confirmation of expectancies to be the critical events in human learning this model may seem more promising than those considered previously However it does not fare so well when applied to the data of Table 4 Using standard derivational techniques (Bush & Mosteller 1955 Ch 8 Estes & Suppes, 1959) one can for example show that the probability of a shift from A_1 on trial n to A_2 on trial $n + 1$ when an E_2 occurs on trial n is equal to the probability of a shift from A_2 on trial n to A_1 on trial $n + 1$ when an E_1 occurs on trial n (both being equal to the quantity $p_n - \alpha_{2,n}$ where $\alpha_{2,n}$ is the mean square of the individual p values on trial n) But from the response proportions in Table 4 we find that the observed conditional proportions of these two types of response shifts are 215 and 129 respectively—a substantial departure from equality Analysis of the other triplets shows that no choice of parameter values enables this model to compare favorably with the other two represented in the table These results tend to support the assumption that under the conditions of this study learning occurs primarily on correction trials as required by a contiguity type association theory rather than on rewarded trials as would be required by theories based on the law of effect

Comparative analyses —The pertinence of the preceding theoretical discussion to our present task lies primarily in the potential usefulness of the type of theoretical model considered for the purposes of comparative analyses of different learning situations Owing to the unevenness of theoretical development in different areas of human learning we have at present only begun to approach a stage in which substantial accomplishment in this direction might be possible In this report we shall be able to include only a couple of brief illustrations of the manner in which comparative analyses might proceed

The combination which appears most nearly ready for fruitful comparisons is that of probability learning and paired associate learning Both versions of stimulus sampling theory the linear model and the pattern model have been applied to the paired associate situation as well as to probability learning The course of events has in fact been quite parallel As with probability learning the linear model was the first member of the family to be tested on paired associate data and the early results appeared surprisingly successful It was no great surprise to find that curves derived from the linear model described paired associate learning curves as well as those derived from Hull s model Thurstone s and others [some comparisons are given in Estes (1961a)] but it proved immediately possible to go further and generate quantitative accounts of such relationships as

that between learning rate and number of response alternatives (Bower, 1962, Estes, 1961a) or the frequency distribution of errors per trial (Estes, 1959a) As in the case of probability learning, however, attempts to account in detail for the fine grain of the data proved more successful with the all or none pattern model than with the linear model (Bower, 1961, 1962, Suppes & Ginsberg, 1962)

Each variety of human learning has its own local problems, and these must be minimized or somehow partialled out in order to make worthwhile comparisons between areas possible In the case of probability learning simplification of conditions seems to have been carried furthest for the two choice predictive situation with noncontingent reinforcement In the case of paired associate learning inter item interference effects probably constitute the major local complication, and therefore comparisons had best be based on data from situations with materials chosen to reduce stimulus and response generalization Subject to these restrictions we have the interesting conclusion, based on studies cited in this and the preceding section, that data arising from these superficially quite different types of experiments can be accounted for in detail by special cases of the same basic model This result I am inclined to take as evidence that the learning process is essentially the same in these two instances For taxonomical purposes, the principal difference between the probability learning and the paired associate learning situations lies in the degree of experimental control over the stimulus patterns available for sampling by the *S* on each trial

Less material is available for comparisons between probability learning and classical conditioning To some previously published analyses of the first few trials of an eyelid conditioning experiment in terms of the pattern model (Estes, 1960b), I can at present add only some fragmentary results from an investigation in progress

As noted earlier, one of the principal local complications in the instance of the eyelid conditioning situation is the response decrement normally observed over a series of massed trials For purposes of a comparative analysis of the acquisition process it seems desirable to begin with data obtained under conditions designed to minimize this source of variation In collaboration with Professor I Gormezano, I have been able to collect a sample of 120 protocols from experiments run in his laboratory with parameters such that curves of CR frequency appear to be approaching unity over a series of 100% reinforced trials [For the conditions under which these data were collected see Gormezano and Moore (1962), Group NV200 or Moore and Gormezano (1961), Control Group] In Table 5 are shown triplet proportions for occurrences and nonoccurrences of CRs in sliding blocks of three trials, pooled over a 70-trial acquisition

TABLE 5

OBSERVED AND THEORETICAL PROPORTIONS OF OCCURRENCES (C) AND NONOCCURRENCES (N) OF CONDITIONED RESPONSES OVER BLOCKS OF THREE SUCCESSIVE TRIALS ($n = 1\ 2 \quad 68$)

Trials			Observed	Theoretical	
n	$n+1$	$n+2$		Linear model	Pattern model
C	C	C	685	685	685
C	C	N	064	064	064
C	N	C	067	070	070
C	N	N	018	015	015
N	C	C	076	078	077
N	C	N	022	020	021
N	N	C	030	0?5	026
N	N	N	037	042	037

series, together with theoretical values calculated from the linear model and the pattern model The equalities across the first line of the table are forced by the estimation procedure, the remaining lines provide a test of the models with 4 degrees of freedom Neither model is far off but of the two, the pattern model comes slightly closer I would emphasize that this example is included only to illustrate a line of investigation that appears worth pursuing further

Complex Learning Processes

There seems to be little room for doubt that it is possible to abstract from the phenomena of probability learning evidence for the pervasive operation of a relatively simple form of conditioning, or associative learning akin to, if not identical with those studied in a number of other standard human learning situations This relatively elementary form of learning appears to depend only upon the successive occurrence of two effective stimuli, a signal and a reinforcing event in experimental parlance Regardless of whether the *S* has been instructed or motivated to learn, the mere succession of events tends to establish (more exactly, has some probability of establishing) a response which in turn depends upon many conditions, of establishing) a response to the signal which may be termed an expectation, anticipation, or preparatory adjustment to the following event In the simplest variants of the verbal conditioning experiment originated by Humphreys the learning, unlearning, and relearning of these expectations as the pairing of a signal with alternative reinforcing events changes from trial to trial seems to be nearly the whole story Depending on instructions, experimental context, and previous experience of the *S*, other forms of behavior, e g, the verbali

zation of hypotheses about causal relationships, can sometimes be observed, especially in the earlier stages of these experiments (e g , Galanter & Smith, 1958), but in the simple predictive situation not involving differential rewards and punishments, these behaviors evidently come and go more or less randomly and play no important role in determining the *S*'s sequence of choices Since the structure of the choice data can be accounted for quite well on the basis of an elementary learning process, it appears that the *S*s' verbalizations in these experiments may be more a matter of talking about what they observe themselves to be doing than of revealing systematic strategies that actually determine their choices to any major extent.

In the simple noncontingent situation, all available cues, whether associated with the experimentally controlled signal, or with past events or responses, have the same correlations with subsequent reinforcing events Consequently, it is not particularly surprising that the data can largely be accounted for by a theory which assumes random sampling of cues by the *S* But as soon as an element of discriminative contingency is introduced, as by making probabilities of reinforcing events depend on properties of the signal (Atkinson, Bogartz, & Turner, 1959, Estes et al , 1957, Popper & Atkinson, 1958) or on preceding events (Anderson, 1960, Engler, 1958, Hake & Hyman, 1953), the situation is quite different In both of these cases some of the new phenomena have been accounted for by natural extensions of the stimulus sampling model, and these results tend to support the parsimonious assumption that the same basic learning processes are operative as in the nondiscriminative situation Especially in the case of event contingencies, however, it has not yet been possible to handle the data in satisfactory quantitative detail The primary source of difficulty here may be the problem of determining the particular portions of the preceding event sequence from which the *S* is sampling cues under any given reinforcement schedule Individual differences in this respect are undoubtedly great, and consequently, as Anderson (1960) has suggested progress toward a more satisfactory interpretation of this experimental variation may come only from intensive study of individual *S*s

Although I find substantial grounds for believing that anticipatory responses, or expectations are learned by an "automatic" associative process in all human learning situations I would not assume that these response tendencies lead directly to overt choices, without intermediary processes, except under especially simplified conditions As soon as rewards and punishments are introduced, the *S* is placed in a conflict situation, in which his overt choices will be determined not only by his predictions, or expectations as to what reinforcing events will ensue on any trial, but also by the approach and avoidance tendencies that have become conditioned to the

various possible trial outcomes. An initial attempt to handle the process of overt response selection under these circumstances within the general framework of stimulus sampling theory has been presented elsewhere (Estes, 1960a, 1962b).

Although I have indicated skepticism as to the likelihood that the learning of higher order response units, or strategies, plays any major role in simple probability learning, I by no means wish to imply that strategies cannot be learned under suitable circumstances. I do have distinct reservations as to the explanatory value of postulating strategies as determiners of overt behavior in situations where we lack direct evidence concerning the existence of such higher order units and also as to the wisdom of leaning heavily on postexperimental interrogation of *Ss* about their mental processes. As has been the case with other aspects of human learning the first steps toward understanding strategies may come from newly designed experiments in which the behaviors in question are recorded as they occur, so that the conditions under which they are learned and the extent to which they control trial to trial responding can be studied directly.

References

Anderson, N. H. An analysis of sequential dependencies. In R. R. Bush & W. K. Estes (Eds.) *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. Pp. 248–264.

Anderson, N. H. Effect of first-order conditional probability in a two-choice learning situation. *J. exp. Psychol.*, 1960, 59, 73–93.

Anderson, N. H., & Grant, D. A. A test of a statistical learning theory model for two-choice behavior with double stimulus events. *J. exp. Psychol.*, 1957, 54, 305–317.

Anderson, N. H., & Grant, D. A. Correction and reanalysis. *J. exp. Psychol.*, 1958, 56, 453–454.

Anderson, N. H., & Whalen, R. E. Likelihood judgments and sequential effects in a two-choice probability learning situation. *J. exp. Psychol.*, 1960, 60, 111–120.

Atkinson, R. C. An analysis of the effect of nonreinforced trials in terms of statistical learning theory. *J. exp. Psychol.*, 1956, 52, 28–32.

Atkinson, R. C., Bogartz, W. H., & Turner, R. N. Supplementary Report: Discrimination learning with probabilistic reinforcement schedules. *J. exp. Psychol.*, 1959, 57, 349–350.

Atkinson, R. C., & Suppes, P. Applications of a Markov model to two person noncooperative games. In R. R. Bush & W. K. Estes (Eds.) *Studies in mathematical learning theory*. Stanford: Stanford Univer. Press, 1959. Pp. 65–75.

Binder, A., & Feldman, S. E. The effects of experimentally controlled experience upon recognition responses. *Psychol. Monogr.*, 1960, 74, Whole No. 496.

Bitterman, M. E. Information and effect in incidental learning. *Amer. J. Psychol.*, 1956, 69, 410–416.

Bower, G. H. Application of a model to paired associate learning. *Psychometrika*, 1961, 26, 255–280.

BOWER, G H An association model for response and training variables in paired associate learning *Psychol Rev*, 1962 69, 34–53
BRAND, H, WOODS P J, & SAKODA, J M Anticipation of reward as a function of partial reinforcement. *J exp Psychol* 1956, 52, 18–22
BRAND, H, SAKODA J M, & WOODS, P J Contingent partial reinforcement and the anticipation of correct alternatives *J exp Psychol*, 1957, 53, 417–424
BRODY, A L Statistical learning theory applied to an instrumental avoidance situation *J exp Psychol* 1957, 54, 240–245
BRODY A L Independence in the learning of two consecutive responses per trial *J exp Psychol* 1958 56, 16–20
BRUNSWIK, E Probability as a determiner of rat behavior *J exp Psychol*, 1939, 25, 175–197
BRUNSWIK E, & HERMA H Probability learning of perceptual cues in the establishment of a weight illusion *J exp Psychol*, 1951, 41, 281–290
BRYANT, S J., & MARICA, J C Strategies and learning models *Psychometrika* 1959, 24, 253–256
BURKE, C J Application of a linear model to two-person interactions In R. R Bush & W K. Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press 1959 Pp 180–203
BURKE C J Some two-person interactions In K J Arrow, S Karlin, & P Suppes (Eds) *Mathematical methods in the social sciences, 1959* Stanford Stanford Univer Press 1960 Pp 242–253
BURKE C J, ESTES, W K., & HELLYER, S Rate of verbal conditioning in relation to stimulus variability *J exp Psychol*, 1954 48, 153–161
BURKE C J, & ESTES, W K. A component model for stimulus variables in discrimination learning *Psychometrika* 1957, 22, 133–145
BUSH R R & MOSTELLER, F *Stochastic models for learning* New York Wiley 1955
BUSH R R., MOSTELLER, F & THOMPSON, G L A formal structure for multiple choice situations In R. M Thrall, C H Coombs & R L Davis (Eds), *Decision processes* New York Wiley, 1954 Pp 99–126
COTTON J W, & RECHTSCHAFFEN, A. Replication report Two- and three-choice verbal conditioning phenomena. *J exp Psychol* 1958 56, 96
DAS J P Mathematical solutions in the acquisition of a verbal CR *J exp Psychol.*, 1961, 61, 376–378
DETAMBEL, M H A test of a model for multiple-choice behavior *J exp Psychol.*, 1955, 49, 97–104
DETAMBEL, M H & STOLUROW L M Probability and work as determiners of multi choice behavior *J exp Psychol.*, 1957, 53, 73–81
EDWARDS W Reward probability amount, and information as determiners of sequential two-alternative decisions *J exp Psychol* 1956 52, 177–188
EDWARDS W Probability learning in 1000 trials *J exp Psychol* 1961, 62, 385–394
ENGLER J Marginal and conditional stimulus and response probabilities in verbal conditioning. *J exp Psychol* 1958, 55, 303–317
ESTES W K. Individual behavior in uncertain situations an interpretation in terms of statistical learning theory In R. M Thrall C H Coombs & R. L Davis (Eds.) *Decision processes* New York Wiley 1954 Pp 127–138
ESTES W K. Theory of elementary predictive behavior In *Mathematical models of human behavior—Proceedings of a symposium* Stamford, Conn. Dunlap & Associates Inc 1955 Pp 63–67
ESTES W K. Of models and men *Amer Psychologist* 1957, 12, 609–617 (a)

ESTES, W K. Theory of learning with constant variable or contingent probabilities of reinforcement *Psychometrika*, 1957, 22, 113–132 (b)

ESTES, W K. The statistical approach to learning theory In S Koch (Ed), *Psychology a study of a science* New York McGraw Hill, 1959 Vol 2 pp 380–491 (a)

ESTES, W K Component and pattern models with Markovian interpretations In R R Bush & W K. Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press 1959 Pp 9–52 (b)

ESTES, W K A random walk model for choice behavior In K J Arrow, S Karlin, & P Suppes (Eds), *Mathematical methods in the social sciences, 1959* Stanford Stanford Univer Press, 1960 Pp 265–276 (a)

ESTES, W K Learning theory and the new mental chemistry *Psychol Rev* 1960 67, 207–223 (b)

ESTES, W K. Growth and function of mathematical models for learning In *Current trends in psychological theory* Pittsburgh University of Pittsburgh Press, 1961 Pp 134–151 (a)

ESTES, W K A descriptive approach to the dynamics of choice behavior *Behavioral Science*, 1961, 6, 177–184 (b)

ESTES, W K Learning theory *Annu Rev Psychol* 1962, 13, 107–144 (a)

ESTES, W K Theoretical treatments of differential reward in multiple choice learning and two-person interactions In J Criswell H Solomon & P Suppes (Eds) *Proceedings of the symposium on mathematical methods in small group processes* Stanford Stanford Univer Press 1962 (b)

ESTES, W K, BURKE, C J, ATKINSON, R C, & FRANKMANN J P Probabilistic discrimination learning *J exp Psychol*, 1957, 54, 233–239

ESTES W K, & JOHNS, M Probability learning with ambiguity in the reinforcing stimulus *Amer J Psychol* 1958, 71, 219–228

ESTES, W K., & STRAUGHAN J H Analysis of a verbal conditioning situation in terms of statistical learning theory *J exp Psychol* 1954, 47, 225–234

ESTES, W K, & SUPPES, P Foundations of linear models In R R Bush & W K Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press, 1959 Pp 137–179

FELDMAN, J On the negative recency hypothesis in the prediction of a series of binary symbols *Amer J Psychol* 1959 72, 597–599

FELDMAN, J, & NEWELL, A A note on a class of probability matching models *Psychometrika* 1961, 26, 333–337

FLOOD, M M On game learning theory and some decision making experiments In R M Thrall C H Coombs & R L Davis (Eds), *Decision processes* New York Wiley, 1954 Pp 139–158 (a)

FLOOD M M Environmental non stationarity in a sequential decision making experiment In R M Thrall, C H Coombs & R L Davis (Eds), *Decision processes* New York Wiley, 1954 Pp 287–300 (b)

FRIEDMAN M P, BURKE, C J COLE, M, ESTES W K KELLER, L, & MILLWARD, R. B Two-choice behavior under extended training with shifting probabilities of reinforcement In R C Atkinson (Ed) *Studies in mathematical psychology* Stanford Stanford Univer Press 1963 Pp 250–291

GALANTER, E H, & SMITH, W A S Some experiments on a simple thought problem *Amer J Psychol* 1958, 71, 359–366

GARDNER, R A Probability learning with two and three choices *Amer J Psychol*, 1957, 70, 174–185

GARDNER, R A Multiple-choice decision behavior *Amer J Psychol*, 1958, 71, 710–717.

GINSBERG, R Three-choice probability learning under correction and noncorrection procedures Paper given at meetings of the Midwestern Psychol Assoc, Chicago, 1959

GOODNOW, J J Determinants of choice distribution in two-choice situations, *Amer J Psychol* 1955, 68, 106–116

GOODNOW, J J, & PETTIGREW, T F Effect of prior patterns of experience upon strategies and learning sets *J exp Psychol*, 1955, 49, 381–389

GOODNOW, J J, & POSTMAN, L Probability learning in a problem solving situation *J exp Psychol*, 1955, 49, 16–22

GOODNOW, J J, RUBINSTEIN, I, & LUBIN, A Response to changing patterns of events *Amer J Psychol*, 1960, 73, 56–67

GORMEZANO, I, & MOORE, J W Effects of instructional set and unconditioned stimulus intensity on the latency, percentage, and form of the eyelid response *J exp Psychol*, 1962, 63, 487–494

GRANT, D A., HAKE, H W, & HORNSETH, J P Acquisition and extinction of a verbal conditioned response with differing percentages of reinforcement *J exp Psychol*, 1951, 42, 1–5

GRANT, D A, HORNSETH, J P, & HAKE, H W The influence of the intertrial interval on the Humphreys' 'random reinforcement' effect during the extinction of a verbal response *J exp Psychol*, 1950, 40, 609–612

GRANT, D A, & SCHIPPER, L. M The acquisition and extinction of conditional eyelid responses as a function of the percentage of fixed ratio random reinforcement. *J exp Psychol*, 1952, 43, 313–320

GREENO, J G Effects of nonreinforced trials in two-choice learning with noncontingent reinforcement. *J exp Psychol*, 1962, 64, 373–379

HAKE, H W, & HYMAN, R. Perception of the statistical structure of a random series of binary symbols *J exp Psychol*, 1953, 45, 64–74

HAKE, H W, GRANT, D A, & HORNSETH, J P Resistance to extinction and the pattern of reinforcement III The effect of trial patterning in verbal "conditioning" *J exp Psychol*, 1951, 41, 221–225

HANANIA, M I Two-alternative learning situations with partial reinforcement *Psychometrika* 1960, 25, 77–90

HAYS, D G, & BUSH R R. A study of group action *Amer Sociol Rev*, 1954, 19, 693–701

HILLIX, W A, & MARX, M H Response strengthening by information and effect in human learning *J exp Psychol*, 1960 60, 97–102

HULL, C L. *Principles of behavior* New York Appleton-Century, 1943

HUMPHREYS, L. G Acquisition and extinction of verbal expectations in a situation analogous to conditioning *J exp Psychol*, 1939, 25, 294–301

JARVIK, M E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols *J exp Psychol* 1951, 41, 291–297

JONES, J E. Deviations from matching behavior in probability learning *J Psychol*, 1961, 52, 335-345

KALMAN, D The effects of intertrial spacing and stimulus variability on verbal conditioning. *J gen Psychol*, 1961, 65, 243–260

KANAREFF, V T, & LANZETTA, J T Effects of task definition and probability of reinforcement upon the acquisition and extinction of imitative responses. *J exp Psychol*, 1960, 60, 340–348

KENT, N Muscle action potentials during verbal conditioning Unpublished PhD dissertation, Indiana Univer, 1958

KOEHLER, J, JR. Role of instruction in two-choice verbal conditioning with contingent partial reinforcement *J exp Psychol*, 1961, 62, 122–125

LABERGE, D Effect of preliminary trials on rate of conditioning in a simple prediction situation *J exp Psychol*, 1959, 57, 20–24 (a)

LABERGE, D A model with neutral elements In R R Bush & W K Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press, 1959 Pp 53–64 (b)

LOTSOF, E J Expectancy for success and certainty of response *Amer J Psychol*, 1959, 72, 600–602

McCORMACK, P D Spatial generalization and probability learning in a five-choice situation *Amer J Psychol*, 1959 72, 135–138

MOORE, J W, & GORMEZANO I Yoked comparisons of instrumental and classical eyelid conditioning *J exp Psychol*, 1961, 62, 552–559

MORSE, E B, & RUNQUIST, W N Probability matching with an unscheduled random sequence *Amer J Psychol*, 1960, 73, 603–607

NEIMARK, E Effects of type of non reinforcement and number of alternative responses in two verbal conditioning situations Unpublished PhD dissertation, Indiana Univer, 1953

NEIMARK, E D Effects of type of non reinforcement and number of available responses in two verbal conditioning situations *J exp Psychol*, 1956, 52, 209–220

NEIMARK, E D, & ROSENBERG, S The effect of 'social' discriminative cues on probability learning *J exp Psychol*, 1959, 58, 302–311

NEIMARK, E D, & SHUFORD E H Comparison of predictions and estimates in a probability learning situation *J exp Psychol*, 1959, 57, 294–298

NICKS, D C Prediction of sequential two choice decisions from event runs *J exp Psychol* 1959, 57, 105–114

OSGOOD, C E, & ANDERSON, L Certain relations among experienced contiguities associative structure, and contingencies in encoded messages *Amer J Psychol* 1957, 70, 411–420

OVERALL, J E A cognitive, probability model for learning *Psychometrika* 1960 25, 159–172

OVERALL, J E, & BROWN, W L A comparison of the decision behavior of rats and of human subjects *Amer J Psychol*, 1959, 72, 258–261

PETERSON, L R Prediction of response in verbal habit hierarchies *J exp Psychol*, 1956, 52, 249–252

POPPER, J, & ATKINSON, R C Discrimination learning in a verbal conditioning situation *J exp Psychol* 1958 56, 21–25

RADLOW, R, & SIEGEL, S Decision making and learning Dept of Psychol, Research Bulletin No 6 (mimeo) State College, Pa Pennsylvania State Univer, 1960

RESTLE, F Theory of selective learning with probable reinforcements *Psychol, Rev*, 1957, 64, 182–191

RESTLE, F *The psychology of judgment and choice* New York Wiley, 1961

ROGERS, O E, WEBB, W B, & GALLAGHER, T J Effect on extinction of restricting information in verbal conditioning *J exp Psychol* 1959, 57, 219–223

RUBINSTEIN, I Some factors in probability matching *J exp Psychol*, 1959 57, 413–416

RUBINSTEIN, I Supplementary report The influence of one stimulus on the prediction of the alternative stimulus in two-choice problems *J exp Psychol* 1961, 62, 311–312

SCHOONARD, J W, & RESTLE, F Analysis of double alternation in terms of stimuli and responses *J exp Psychol*, 1961, 61, 365–367

SHELLY, M W II Effects of increments of reinforcement in human probability learning *J exp Psychol*, 1960, 59, 345–350 (a)

SHELLY, M W II Supplementary report Further factors affecting the probability of changing responses *J exp Psychol*, 1960, 59, 430–431. (b)

SIEGEL, S Decision making and learning under varying conditions of reinforcement. *Annals N Y Acad Sci*, 1961, 89, 766–782

SIEGEL, S., & GOLDSTEIN, D A Decision making behavior in a two-choice uncertain outcome situation, *J exp Psychol*, 1959, 57, 37–42

SIMON, H A. A comparison of game theory and learning theory *Psychometrika*, 1956, 21, 267–272

SOLLEY, C M, & MESSICK, S J Probability learning the statistical structure of con cepts, and the measurement of meaning *Amer J Psychol*, 1957, 70, 161–173

STERNBERG, S H A path-dependent linear model In R. R. Bush & W K Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press, 1959 Pp 308–339 (a)

STERNBERG, S H Application of four models to sequential dependence in human learning In R. R Bush & W K Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press, 1959, Pp 340–380 (b)

STEVENSON, H W, & ZIGLER, E F Probability learning in children *J exp Psychol*, 1958, 56, 185–192

STRAUGHAN, J H Human escape learning in relation to reinforcement variables and intertrial conditions *J exp Psychol* 1956, 52, 1–8

SUPPES, P, & ATKINSON, R. C *Markov learning models for multiperson interactions* Stanford Stanford Univer Press, 1960

SUPPES, P, & FRANKMANN, R W Test of a stimulus sampling theory for a con tinuum of responses with unimodal noncontingent determinate reinforcement. *J exp Psychol*, 1961, 61, 122–132

SUPPES P, & GINSBERG, R Application of a stimulus sampling model to children s concept formation of binary numbers with and without an overt correction response *J exp Psychol*, 1962, 63, 330–336

TAUB, H A, & MYERS, J L Differential monetary gains in a two-choice situation *J exp Psychol*, 1961, 61, 157–162

VOSS, J F, THOMPSON, C P, & KEEGAN, J H Acquisition of probabilistic paired associates as a function of S R₁ S R₂ probability *J exp Psychol*, 1959, 58, 390–399

WOODS, P J The relationship between probability difference (π_1-π_2) and learning rate in a contingent partial reinforcement situation *J exp Psychol*, 1959, 58, 27–30 (a)

WOODS, P J The effects of motivation and probability of reward on two-choice learning *J exp Psychol* 1959, 57, 380–385 (b)

An Evaluation of Stimulus Sampling Theory

COMMENTS ON PROFESSOR ESTES' PAPER[1]

NORMAN H. ANDERSON

University of California, Los Angeles

The evaluation of any theory has two aspects, one concerning its present state, the other having regard to its possible elaboration and development. These two aspects present obviously unequal difficulty to a discussant, but both are needed for a balanced view.

Although Professor Estes appears well satisfied with the present state of stimulus sampling theory, my own assessment of the existing evidence leads me to a different conclusion. With respect to the future of the theory, however, I think that the situation is more hopeful. The following comments will take up three or four of the main areas of investigation which appear to have the most importance in understanding these two aspects of the theory.

Since the pattern model used by Estes is relatively unfamiliar, it will be useful to address a few remarks to it before proceeding. At least roughly, the pattern model may be considered as a special case of the original stimulus element sampling model with the single restriction that exactly one element (pattern) be sampled each trial. Conditioning can occur only on incorrect trials since a correct response means that the single sampled pattern was already conditioned. Conditioning of the sampled pattern occurs in an all or none manner as in the stimulus element model. In the special case of the pattern model in which only one pattern exists ($N = 1$), learning is also all-or-none. However, if more than one pattern exists, as would be required by the eyelid conditioning data of Estes' present Table 5 for instance, then learning occurs in an incremental manner. With certain exceptions, the pattern model and the linear operator model make the same predictions, as Estes has noted, and it will not be necessary to consider them separately in what follows.

[1] This work was supported by National Science Foundation grant, G 12986.

Reinforcement Schedules and Mean Response

As is well known, much of the interest in stimulus sampling theory arose from the predictions of "matching" behavior in the two-choice, Humphreys' task (Estes & Burke, 1953, Estes & Straughan, 1954), and much of the support for the theory resides in the generally good agreement found in numerous studies since the initial matching observations of Grant and Hake (1949) Of course, as Estes has noted, nonmatching does occur under certain conditions and indeed much of the work critical of the theory has been oriented toward demonstrating this fact However, it seems to me that this matching controversy has not been too useful because, for various reasons, tests of the matching prediction have somewhat limited theoretical relevance

As Estes has observed, stimulus sampling theory is a general rather than a situational theory Application of a general theory to a given situation usually rests on certain simplifying assumptions to the neglect of certain variables Perfect quantitative predictions are not to be expected because of the possible influence of the neglected variables The point may be illustrated in a somewhat different way with the Burke and Estes (1957) generalization of the original Estes and Straughan (1954) model This generalized model attempts to take trace stimuli into account, which seems quite reasonable from an S—R point of view, and may be considered as a natural state in the development of the theory However, this generalized model does not predict matching although, to be sure, the deviations are not large

That tests of mathematical models are usually best interpreted within a larger theoretical context has been elaborated by Grant (1962) and there is no doubt about the sensibleness of this view From this standpoint, however, the predictive precision of particular models can be misleading since the general theory usually does not have the same degree of precision

The extent to which models based on different assumptions tend to agree in predictions of mean asymptotic response level should be kept in mind The linear operator and pattern models discussed by Estes both predict matching although in the former, reinforcement is independent of the overt response, whereas in the latter, reinforcement occurs only on incorrect trials Moreover, these two models are but the extremes of a family of models (Estes, 1959), all of which predict matching Feldman and Newell (1961), working from a considerably different rationale, have also exhibited a large class of matching models

There would seem to be some natural tendency for any simplified model which incorporates the empirical fact that the stimulus events act as reinforcers to predict either matching or something close to it for the standard

two-choice task This may stem in part from the symmetry of the task and because of the fact that matching almost necessarily occurs for the 100 0, 50 50, and 0 100 reinforcement schedules The task thus imposes somewhat greater limits on the range of behavior than would be the case with eyelid conditioning, for instanee, for whicb 50% reinforcement does not necessarily imply 50% response

Possibly some of the interest in mean asymptotic response under the two choice frequency schedule has arisen from a general disbelief that Ss act the way they do A priori one might expect Ss to act "rationally,' predict ing always the more frequent event, and indeed some Ss do However, the models do not predict they will not, they simply do not apply

Taken as a whole, the preceding comments suggest that mean asymptotic response in the two-choice frequency scedule is of limited theoretical use fulness In addition, even where deviations from prediction are seriously large, the mean response rate may not be very informative about the eause of the diserepaney Aceordingly, it might be expected that other descriptive statistics or other experimental situations could yield more cogent evidenee about the adequacy of the theory

This diseounting of mean asymptotic response should be taken with due caution sinee future developments may put the matter in a different light It should also be noted that the deviations from matching often found in animal work (Bitterman Wodinsky, & Candland 1958 Wilson 1960) may deserve more serious attention since in these experiments the asymptotic response is presumably less complicated by pre experimental habits than is the case with human Ss However, I think that the above comments represent the preseot state of affairs in human experimentation fairly well

In his summary of evidence, Estes notes that results from reinforcement schedules other than the one just discussed involving for instance con tingent reinforcement, or the use of more than two choices have yielded generally poorer agreement with predictions of asymptotic response Ap parently, the discrepancies are considered large enough to warrant some theoretical concern, but it may be possible to resolve them with further work Estes notes that the instructioos may be the crucial variable in con tingent reinforcement, and some further evideoce for this suggestion comes from the generally good results of the Suppes and Atkinson (1960) work on bicontingent, two person interaction tasks The use of more than two choices probably deserves further study because of the evident discrepancy with the two choice results However, Atkiosons (1961) recent generaliza tion of stimulus sampling theory shows some promise of handling the non matching in the three-choice task found by Gardner (1957), although Atkinson s model may predict noomatching in the two-choice task as well

CONDITIONAL PROBABILITY SCHEDULES

There is however, a two-choice noncontingent schedule which deserves more detailed comment because of its direct theoretical relevance This is the conditional probability schedule introduced by Hake and Hyman (1953) For convenience, this schedule will be called the conditional schedule, and the standard two-choice, noncontingent task in which event frequency is varied, will be called the frequency schedule

In conditional schedules, the probability of an event on one trial depends on which event occurred on the previous trial A conditional probability of 0 corresponds to a pure single alternation sequence, RLRL ,whereas a conditional probability of 1 corresponds to a pure repetition sequence, RRRR Conditional probability may have any value between 0 and 1, and will correspond to greater or lesser degrees of alternation and repetition tendency within the sequence of stimulus events

The immediate question is whether the linear operator and pattern models discussed by Estes are to be considered applicable to the conditional schedule If so, their validity would be open to serious question For instance, both models predict that performance on a single alternation schedule will be poorer than chance whereas, of course, it is perfect

Estes indicates that the models are not to be considered as applicable to conditional schedules because the differential correlations between the present event and various combinations of past events and responses give rise to discriminative contingencies In principle at least, it would be possible to treat the conditional schedules by taking account of such contingencies Thus, for the single alternation sequence, the event on one trial would be a cue to predict the opposite event on the next trial Of course, this argument rests on the assumption that *S*s have memory traces for past events, but of this there is no doubt since otherwise they could not learn, for instance, the single alternation schedule, nor would they show a negative recency effect More generally, Hake and Hyman (1953) concluded that their results showed that *S*s respond in terms of the preceding events and responses, and this conclusion has been supported by the work of numerous other investigators

If subjects have memory traces of past events when they receive a conditional schedule, then it can hardly be doubted that they have memory traces of past events when they receive a frequency schedule But neither the linear operator model nor the pattern model allow memory traces of past events Memory traces are allowed in the general theory, of course, but are ruled out by the sampling assumptions of the two models Both models assume that all stimulus elements or patterns have equal, or at

least trial independent, sampling probabilities If memory traces were allowed, the sampling probabilities would vary from trial to trial depending on what previous events and responses happened to occur

We thus face a theoretical inconsistency If the theory requires trace stimuli to handle the conditional probability schedule, it cannot properly disallow them when it treats the frequency schedule

According to the theory, trace stimuli present on any trial will automatically become conditioned in the same way as any other stimuli A RLRL single alternation subsequence occurring by chance in a frequency schedule will develop conditioned traces in essentially the same way as if it had occurred in a conditional schedule However, the trace stimuli present on any trial not only depend on the past events and responses but also influence the present response on the given trial For this reason response probability will fluctuate from trial to trial in a way not described by either model Consequently, if the theory is correct, then neither model can account for the trial to-trial dependencies in the behavior unless, of course, only a negligible amount of trace stimulation is present

It could still be hoped that these trial to trial discrepancies would cancel out over a number of trials so that the trace stimuli would have no net effect on the mean asymptotic response in the frequency schedule Some evidence that this is true in point of empirical fact is seen in the results of Hake and Hyman (1953) and of Engler (1958) who used schedules which incorporated joint variation in frequency and in conditional probability From the observed response to the conditional aspect of the sequences, it is known that trace stimulation was present and active However, this had no apparent effect on the response to the frequency aspect of the schedules In the Hake and Hyman study, two 75 25 frequency schedules which embodied different conditional probabilities yielded the same asymptotic mean response The more extensive data of Engler appear to be in essential agreement with this result There is thus some reason to believe that the effect of at least certain types of trace stimulation does in fact cancel out to leave no net effect on the mean asymptotic response in the frequency schedule It may be noted, incidentally, that the Hake and Hyman technique of joint variation of frequency and conditional probability would appear to be a useful tool for investigating the nature and effect of trace stimuli

Unfortunately, the cited empirical evidence is not to the point of the present problem which is one of theory, namely, whether the theory would make the same predictions as the current models do if trace stimulation were taken into account Estes comments in this connection that in a frequency schedule, all combinations of past events and responses have the same correlation with subsequent reinforcing events Although this means

that the *S*s cannot gain any mathematical advantage by responding in terms of various patterns of past events and responses, the theoretical implications of this comment are not obvious If trace stimuli are present, the theory postulates that they are conditioned in the same way as other stimuli, regardless of whether this is to the *S*'s advantage or not That taking account of trace stimuli would change the predictions of mean asymptotic response is adumbrated by the Burke and Estes (1957) development, which represents one way of taking account of trace stimuli from the preceding trial, but which does not predict matching

It may be possible to get a consistent development of the theory that gives a realistic account of trace stimuli If so, the resultant predictions may be fairly close to those made by the present models This would seem rather likely for mean asymptotic response in the two choice frequency schedule, but rather unlikely as far as trial to trial changes in the behavior are concerned If this is correct, applications of the present models must be interpreted with appropriate caution since in the last analysis, these models are based directly on the trial to trial reinforcement effects

Sequential Dependencies

Sequential dependencies refer to the dependence between the response on one trial and the events and responses of the preceding trials As has been emphasized on a number of occasions (Anderson & Grant, 1957, 1958, Anderson, 1959, 1960), sequential dependencies are generally more useful in the study of learning than is the mean response curve These dependencies pick up the trial to trial increments and decrements in response and so give the most direct measure of the action of the reinforcing stimuli

Sequential dependencies which include reference to previous responses are often difficult to interpret because of a probability selection effect, and because of an individual difference selection effect (Anderson, 1959, 1960) Evidence from such dependencies must be considered equivocal without a more serious study of individual differences than has yet been made Accordingly, only stimulus dependencies which involve reference to just the preceding stimuli will be considered here

Perhaps the best known sequential effects are the positive and negative recency effects (Jarvik 1951) As Jarvik notes, the positive recency effect means essentially that the presumptive reinforcing stimuli are reinforcers in fact, or in other words, that learning occurs This effect has not attracted much specific comment

The negative recency effect, first observed by Humphreys (1939), has excited considerably more interest because it is inconsistent, not only with

There are a number of sequential dependencies other than those already mentioned and it is appropriate to exhibit results from a fairly extensive empirical investigation (Anderson, 1960) The data most relevant to present purposes are from the 40 *Ss* of Group 50 which received a 50 50 frequency schedule With certain changes in notation, these dependencies are reproduced in Table 2 The theoretical expressions at the foot of each column give the predictions derived from the linear operator model (Anderson, 1959), and the pattern model makes the same predictions

Each entry in the table is a difference of two sequential dependencies In the first two columns, each term of the difference is based on 2000 instances, the corresponding figure for each entry in the last four columns is 500 It thus needs no formal statistical analysis to suspect some serious shortcomings in the models If these discrepancies are to be attributed to trace stimuli, then it is evident that these stimuli play a nonnegligible role In any case, I cannot agree with Estes' general statement about frequency schedules that " the structure of the choice data can be accounted for quite well on the basis of an elementary learning process " (p 122)

One technical comment on these data should perhaps be added The stimulus sequences were balanced over blocks of 50 trials so that, since the model predictions assume completely random sequences, it may be questioned whether they are applicable However, the randomness assumption is not part of the models but is used simply for convenience in making the derivations Since the balancing restriction was mild, it would not be expected to have any noticeable effect on the derivations It is possible to check this using a stat rat procedure and this was done for the linear operator model with the expected results From the standpoint of a trace model, it would again be expected that the balancing restriction would have a negligible effect as long as the memory traces extended only over a few preceding trials However, if memory traces extend over a large number of preceding trials, or if there were other long range effects, the situation might be different

The Friedman, Burke, Cole, Estes, and Millward Experiment

The model predictions for the response dependencies from the Friedman et al. (1960) experiment which Estes has shown appear quite impressive Although a complete evaluation must await a fuller report of the data, and method of parameter estimation, this experiment merits some comment because of its potential theoretical and methodological significance

In his discussion of this experiment, Estes has suggested that such theo-

TABLE 2

SEQUENTIAL DEPENDENCIES FOR TWO-CHOICE, NONCONTINGENT, PROBABILITY LEARNING DATA

Trials	Subsequences of Preceding Reinforcing Stimuli							
	$E_1 - E_2$	$E_1E_1 - E_2E_1$	$E_1E_1E_1 - E_2E_1E_1$	$E_1E_2E_1 - E_2E_2E_1$	$E_1E_1E_1E_1 - E_2E_1E_1E_1$	$E_1E_2E_2E_1 - E_2E_2E_2E_1$	$E_1E_1E_2E_1 - E_2E_1E_2E_1$	$E_1E_2E_1E_1 - E_2E_2E_1E_1$
1–100	14	−01	−16	−16	−10	00	11	03
101–200	22	12	−01	−13	−06	−01	11	−01
201–300	24	04	−04	−13	−12	−03	10	−02
301–400	24	09	−03	−18	−04	05	15	01
401–500	20	03	01	−11	01	00	03	03
101–500								
Observed	22	07	−02	−14	−05	00	10	00
Predicted	—	17	14	14	10	10	10	10
Theoretical	θ	$\theta(1 - \theta)$	$\theta(1 - \theta)^2$	$\theta(1 - \theta)^2$	$\theta(1 - \theta)^3$	$\theta(1 - \theta)^3$	$\theta(1 - \theta)^3$	$\theta(1 - \theta)^3$

NOTE The predicted values are given for the data averaged over Trials 101–500, using the listed theoretical expressions together with the estimate of θ given by the first data column

Entries are differences of two sequential dependencies, decimal points omitted defined by the stimulus subsequences heading each column For instance the second data column lists the proportion of A_1 responses given that the two preceding reinforcing stimuli were E_1E_1, minus the proportion of A_1 responses given that the two preceding reinforcing stimuli were E_2E_1 Data are for Group 50, transcribed with necessary sign changes from Tables 2, 3, 4, and 6 of Anderson (1960)

retical discrepancies as have heretofore been found can be eliminated by the use of practiced *Ss* It might be expected that whatever pre-experimental response tendencies *Ss* bring to the task would tend to disappear over trials as often happens, for instance, with the negative recency effect One could thus hope that extended training would wash out such tendencies, revealing in pure form the elementary S—R conditioning process envisaged in the theory This is an attractive proposition and, if correct, it would be of great theoretical importance

A somewhat different point of view is also possible Instead of assuming that practice trains out the interfering response tendencies, it may be suggested that practice trains in the very behavior described by the models, that different practice regimes will train in different sorts of behavior, and that each will require a different model

To see how this second point of view would apply to the experiment in question, it should first be noted that the data that Estes reports are from Trials 97–192 of the 80 20 schedule given on Day 3 However, over the first two days of the experiment, sixteen 48 trial blocks were given, each successive block based on a different one of the nine frequency schedules ranging from 10 90 to 90 10 with alternate blocks on the 50 50 schedule In the training sequences, therefore, an event occurring on one trial will have a greater than chance probability of occurring on the next trial, even though each event occurs equally often on the average Consequently, the first order as well as the higher order conditional probabilities in the training sequence are greater than chance, and the sequence may be roughly characterized by its tendency toward unusually long runs on the two events It is thus evident that there is an element of discriminative contingency involved

From a common sense standpoint, certainly, one might wonder whether *Ss* would not be affected by the tendency toward repetition in the stimulus sequence, to say nothing of the shifts in stimulus frequency from block to block, with appropriate changes in their behavior If so, they could become rather dependent on the previous stimulus, developing a tendency toward *repetition responding*, that is, predicting next that event which occurred last. In terms of the responses, this would mean that a correct response would tend to be repeated, whereas an incorrect response would tend to be changed Precisely this sort of behavior is envisaged by the pattern model

Although this common sense argument has no evidential status, it does set out the two lines of evidence needed to buttress the view that the terminal behavior was specifically trained in, and it also indicates where this evidence is to be sought. These lines of evidence are, first, that different initial training schedules produce different sets or modes of responding,

second, and more important, that these different sets are reasonably persistent, lasting considerably beyond their establishment

Evidence along the first line is to be seen in the results of Goodnow and Pettigrew (1955) These investigators used both Short-Run" and Long-Run" 50 50 sequences and found that they yielded different response tendencies In particular, the behavior under the Long Run sequences, constructed in a manner quite similar to those used by Friedman et al, was interpreted as exhibiting a 'Win-Stay, Lose Shift 'strategy" This strategy may evidently be considered as the cognitive counterpart of the pattern model

The results of Anderson (1960) support those of Goodnow and Pettigrew, and also give a firm foundation to the second line of evidence The high and low conditional probability sequences of this experiment produced different acquisition behavior and, more importantly, this difference was maintained at a high level over several hundred transfer trials on a 50 50 random sequence common to all conditions In addition, this difference was maintained over a 2- to 6-day rest which even produced a certain amount of spontaneous recovery It would thus appear that the nature of the terminal behavior after extended training will depend quite strongly on the type of training which has been given

This conclusion bears directly on the theoretical interpretation of the Friedman et al experiment The high conditional probability schedules just mentioned are similar to those used by Friedman et al in having a tendency for long runs on one stimulus to alternate with long runs on the other stimulus The use of these high conditional probability schedules in training produced a rather high rate of about 70% repetition response in the 50 50 transfer sequence As has been noted, this is the sort of behavior which the pattern model is well fitted to describe

At the same time, however, the low conditional probability schedules produced in transfer a markedly lower repetition response rate, slightly under 50% to be exact, and this sort of behavior is poorly described by the pattern model Taken as a whole, then, the data indicate that different training induces different response tendencies of considerable persistence, and that different models will be required to describe the terminal behavior following different training regimes

No doubt it would be desirable if extended practice simplified the behavior sufficiently so that some model described it adequately, even if the model were specific to the particular practice conditions Although this might represent mainly a descriptive success, its value would probably be sufficient to make use of practiced *Ss* quite worthwhile

It is more difficult to assess the theoretical implications of a need for a multiplicity of models for a given task The problem would be whether the

various models could be unified in terms of a general theory On this question, it seems hazardous to make any forecast except to note that it may be quite difficult to account for persistent differential transfer effects of the type noted above

Stimulus Variation Experiments

A final class of experiments bearing on stimulus sampling theory comes from situations in which there is an active manipulation of the signal stimulus These include the Burke, Estes, and Hellyer (1954) study of stimulus variability, and various experiments on probabilistic discrimination learning (e g , Atkinson, Bogartz, & Turner, 1959, Estes & Burke, 1957, Estes, Burke, Atkinson, & Frankmann, 1957, Schoeffler, 1954) The results of these experiments have generally been in at least qualitative agreement with the models employed and in certain cases the models fit extremely well However, it should be noted that the imperfect discrimination in some of these experiments requires nonzero stimulus overlap, whereas the perfect discrimination in other experiments requires zero stimulus overlap (Lawrence, 1958) This difference reflects a theoretical inconsistency which is not yet resolved although Atkinson's (1960) observing response formulation of discrimination learning appears to offer very great promise

The main characteristics which distinguish stimulus sampling theory from other reinforcement theories are its representation of the stimulus, and of the stimulus response connection For this reason, it would seem that experiments of the class under discussion, in which there is an experimental manipulation of the discriminative stimuli, afford the best ground for theoretical exploration

Stimulus Sampling Theory

I think it is fair to say that the principal accomplishment of stimulus sampling theory as applied to probability learning has been the emphasis it has placed on quantifying the action of the reinforcing stimuli Compared to the assorted normative and static theories which have been popular in the study of probability learning the mathematical learning model approach has several advantages as Estes has noted In the first place, it begins with the effects of the reinforcing stimuli This constitutes one of the basic aspects of the behavior, an aspect which is neglected by the static theories Moreover, it affords a uniform method of attack on a wide variety of reinforcement schedules Finally, it yields precise predictions for a variety of data statistics which yield a degree of testability enjoyed

by few other approaches These advantages are well known, of course, but I do not think that they can be valued too highly

At the same time, there are two negative aspects to be taken into account, one having to do with the behavior, the other pertaining to the theory

As has been noted, the existing data indicate that there is much in the behavior which is neglected by the stimulus sampling models Indeed, it would almost seem that what is neglected is just what is most of interest, and in this respect one must question the net usefulness of the theory Although the theory has certainly stimulated a large volume of work by many highly competent investigators over the past decade, the amount of knowledge gained about the behavior appears quite meager and this seems to me to stem to a considerable extent from preoccupation with the theory Thus, for example, much effort has been spent on the matching controversy whereas the early work of Jarvik (1951) and of Hake and Hyman (1953) has until recently been largely ignored although this work is noteworthy not only for its empirical and methodological significance, but also for the serious theoretical problems which it raises For probability learning itself, it would seem that more rapid progress would be got from a more empirical attack coupled perhaps with the frankly descriptive learning model approach of Bush and Mosteller (1955)

Presumably, however, stimulus sampling theory is not primarily concerned with probability learning behavior for its own sake, but instead uses the experimental situation as a tool for testing and developing the theory For this purpose, the question of how large a fraction of the behavior the models predict is not always directly to the point The question is rather to what degree the applications of the models support the basic conceptualizations of the theory

It is in this latter respect that I feel uneasy about the present state of affairs The failure to account for the sequential dependencies shows that the model assumptions about the action of the reinforcing stimuli are inadequate More generally, the present models do not allow for any memory effect and yet the existing evidence gives every indication that memory, in one role or another, plays an important part In addition, there is also some evidence of a transfer effect which is outside the scope of the models and perhaps of the theory itself

It is possible, of course, that these memory and transfer effects are merely complicating factors which have only to be added on to the present theoretical framework However, it is also possible that these effects reflect the main behavior characteristics of which the present models represent only some ill defined average Judgment on this issue is difficult, but it

would seem that until the theoretical models are developed to a more realistic level, the present applications must be viewed with some reserve.

On the face of it, there would seem to be no reason that the theory could not be extended to handle the problem of trace stimuli. There might be some technical difficulties in achieving a consistent theoretical treatment, of course, but one of the attractions of the theory is that it seems ideally suited to represent memory traces in terms of stimulus elements or patterns. However, different traces would probably have different sampling probabilities and these would necessarily vary from one trial to the next. The present simple models would no longer apply and the mathematical picture would rapidly become complicated. Because of this, and because there would be no definite limit to the number of possible stimulus elements or patterns, it might be difficult to get unambiguous tests of the theory.

No psychologist was ever at a loss to criticize a theory, and in doing so it is perhaps too easy to slight past accomplishments, and to magnify present difficulties and future problems. On the whole, however, it seems to me that the ordinary probability learning situation is not the most useful area in which to develop the theory. I have been partial to stimulus sampling theory because of its clean simplicity, but this is likely to be lost without a greater control of the effective stimulus than seems likely to be obtained in the probability learning task. Elaboration of the theoretical models may very well enjoy considerable success, but such success may not reflect very directly on the basic theoretical conceptions of the stimulus, and the stimulus response connection.

Animals or children may yield more useful probability learning data than has been obtained from adult humans, and the extended training technique of Friedman et al. certainly deserves further exploration. Replacement of the guessing by a learning aspect as in the probabilistic paired-associate work of Voss, Thompson, and Keegan (1959) might also be helpful. There is also, of course, a wide range of learning tasks outside the probability learning area, and Estes refers to some initial applications of the theory to such tasks. It is perhaps from such work that the most rapid progress is to be expected. The mathematical tools and sophistication already developed should be of considerable assistance in this work.

References

ANDERSON, N. H. An analysis of sequential dependencies. In R. R. Bush & W. K. Estes (Eds.) *Studies in mathematical learning theory.* Stanford: Stanford Univer. Press, 1959. Pp. 248–264.

ANDERSON, N. H. Effect of first-order conditional probability in a two-choice learning situation. *J. exp. Psychol.*, 1960, 59, 73–93.

ANDERSON, N H, & GRANT, D A A test of a statistical learning theory model for two-choice behavior with double stimulus events *J exp Psychol* 1957, 54, 305–317

ANDERSON, N H, & GRANT, D A Correction and reanalysis *J exp Psychol*, 1958, 56, 453–454

ANDERSON, N. H, & WHALEN, R E Likelihood judgments and sequential effects in a two-choice probability learning situation *J exp Psychol*, 1960, 60, 111–120

ATKINSON, R C. A theory of stimulus discrimination learning In K J Arrow, S Karlin, & P Suppes (Eds), *Mathematical methods in the social sciences* Stanford Stanford Univer Press, 1960 Pp 221–241

ATKINSON, R. C A generalization of stimulus sampling theory *Psychometrika* 1961, 26, 281–290

ATKINSON, R C, BOGARTZ, W H, & TURNER, R N Supplementary report. Discrimination learning with probabilistic reinforcement schedules *J exp Psychol*, 1959, 57, 349–350

BITTERMAN, M E, WODINSKY, J, & CANDLAND, D K Some comparative psychology *Amer J Psychol*, 1958, 71, 94–110

BURKE, C J, & ESTES, W K A component model for stimulus variables in discrimination learning *Psychometrika*, 1957, 22, 133–145

BURKE, C J, ESTES, W K, & HELLYER, S Rate of verbal conditioning in relation to stimulus variability *J exp Psychol*, 1954, 48, 153–161

BUSH, R R, & MOSTELLER, F *Stochastic models for learning* New York Wiley, 1955

ENGLER, J Marginal and conditional stimulus and response probabilities in verbal conditioning *J exp Psychol*, 1958, 55, 303–317

ESTES, W K. Component and pattern models with Markovian interpretations In R. R Bush & W K Estes (Eds), *Studies in mathematical learning theory* Stanford Stanford Univer Press, 1959 Pp 9–52

ESTES, W K., & BURKE, C J A theory of stimulus variability in learning *Psychol Rev* 1953, 60, 276–286

ESTES, W K, & BURKE, C J Application of a statistical model to simple discrimination learning in human subjects *J exp Psychol*, 1955, 50, 81–88

ESTES, W K, BURKE, C J, ATKINSON, R. C., & FRANKMANN, J P Probabilistic discrimination learning *J exp Psychol*, 1957, 54, 233–239

ESTES W K, & STRAUGHAN, J H Analysis of a verbal conditioning situation in terms of statistical learning theory *J exp Psychol* 1954, 47, 225–234

FELDMAN, J, & NEWELL, A A note on a class of probability matching models *Psychometrika* 1961, 26, 333–337

FRIEDMAN, M P, BURKE, C J, COLE, M, ESTES, W K, & MILLWARD, R B Extended training in a non contingent two-choice situation with shifting reinforcement probabilities Paper given at meetings of the Psychonomic Society, Chicago, 1960

GARDNER, R A Probability learning with two and three choices *Amer J Psychol*, 1957, 70, 174–185

GOODNOW, J J, & PETTIGREW, T F Effect of prior patterns of experience upon strategies and learning sets *J exp Psychol* 1955, 49, 381–389

GRANT, D A Testing the null hypothesis and the strategy and tactics of investigating theoretical models. *Psychol Rev*, 1962, 69, 54–61

GRANT, D A, & HAKE, H W Acquisition and extinction of the Humphreys' verbal response with differing percentages of reinforcement" *Amer Psychologist*, 1949, 4, 226 (Abstract)

HAKE, H W & HYMAN R. Perception of the statistical structure of a random series of binary symbols *J exp Psychol*, 1953, 45, 64–74

HUMPHREYS, L. G Acquisition and extinction of verbal expectations in a situation analogous to conditioning *J exp Psychol* 1939, 25, 294–301

JARVIK, M E Probability learning and a negative recency effect in the serial anticipation of alternative symbols *J exp Psychol.*, 1951, 41, 291–297

LAWRENCE, D H Learning. *Ann Rev Psychol*, 1958, 9, 157–188

SCHOEFFLER, M S Probability of response to compounds of discriminated stimuli *J exp Psychol*, 1954, 48, 323–329

SUPPES P & ATKINSON R. C *Markov learning models for multiperson interactions* Stanford Stanford Univer Press 1960

VOSS, J F, THOMPSON, C P, & KEEGAN, J H Acquisition of probabilistic paired associates as a function of S-R, S R, probability *J exp Psychol*, 1959, 58, 390–399

WILSON, W A, Jr Supplementary report Two-choice behavior of monkeys *J exp Psychol*, 1960, 59, 207–208

Short-Term Memory and Incidental Learning[1]

LEO POSTMAN
University of California, Berkeley

This paper reviews experimental methods and findings in the study of short-term retention and incidental learning. There are important continuities between the theoretical and methodological problems in these two areas. Both types of studies are concerned with basic capacities and dispositions which the learner brings to the experimental situation and which determine the initial reception and immediate storage of information. Such experiments are not primarily designed to investigate the laws governing the integration of responses and the growth of associative strength. Rather, they are concerned with a detailed analysis of some of the conditions which limit and bias the subject's (*S*'s) responses in a learning situation. Extended practice must take its departure from these initial dispositions of the learner. In that sense the classes of experiments considered here and the designs for the study of rote learning and forgetting supplement each other.

For purposes of delimiting our area of inquiry, it is convenient to begin with a classification of experiments on the basis of the temporal arrangement of exposures to the learning materials and of tests of performance. In the conventional rote learning experiment, exposures and tests follow each other in a fixed order within successive trials. When a single presentation of the learning materials is followed without delay by a test of performance, the measures of retention define the amount of immediate memory. In studies of short-term retention the same experimental arrangements are typically used except that the interval between exposure and test is varied systematically. The delimitation of "short" in the phrase, short-term retention, is necessarily arbitrary. In actual practice, retention intervals of the order of seconds or minutes are used. The intensive exploration of such very brief retention intervals has usually been undertaken with a view to gauging the rate of decay of the stimulus trace. Since a test of immediate memory or short-term retention records the *S*'s initial response to a set of learning materials, it may also be expected to yield information about pre-experimental habits which become sources of transfer in the mastery of

[1] The preparation of this paper was facilitated by a grant from the National Science Foundation to the Institute of Human Learning, University of California, Berkeley, and by Contract Nonr 222(90) with the Office of Naval Research.

new tasks Foremost among these sources of transfer are the *S*'s modes of classifying and grouping the stimulus items

In a given experimental situation the presentation of the materials to the *S* requires not only a controlled physical arrangement but also the choice of an instruction stimulus In the study of sensory and perceptual processes it is explicitly recognized that a psychophysical function is determined jointly by the conditions of physical stimulation and the instruction stimulus (Graham, 1952) Similarly, in a learning experiment the instruction stimulus, which directs the *S* to the relevant characteristics of the materials and specifies the responses required of him, interacts with the conditions of presentation to determine the course of acquisition The distinction between intentional and incidental learning refers to the degree to which the instruction stimulus prepares the *S* for the test of performance Although the difference between the two kinds of learning is usually stated in an all-or none fashion, there is every reason to believe that set to learn may vary continuously in degree as a function of the instruction stimulus Thus the terms, intentional and incidental, define the extremes of a dimension rather than a dichotomy

The systematic significance of studies of incidental learning lies primarily in the opportunity which they afford to identify the principles governing selective discrimination and retention of environmental events Instructions to learn delimit the characteristics of the stimuli which are relevant to the *S*'s task and mobilize instrumental acts such as rehearsal which maximize the amount of retention In the absence of such instructions the *S*'s dispositions to respond selectively to some characteristics of the stimuli rather than to others come into play, as do his habits of classifying and grouping the stimuli The effects of such differential responses on retention can be assessed best when deliberate rehearsal is minimized In the interest of this objective, studies of incidental learning typically use a single, or at most, a few, presentations of the materials and short retention intervals Thus, the experimental arrangements are characteristically those for the measurement of short term retention

In spite of these continuities between studies of short term retention and incidental learning, it will be convenient to maintain a distinction between these two areas of investigation for purposes of the present discussion A separate treatment will bring out the differences in theoretical concerns and methodological issues without losing sight of the continuities

Short Term Retention

The problems considered in experimental studies of immediate memory and short term retention fall into three related categories The first concerns

the limitations on our capacity for immediate reproduction of the materials to which we have been exposed What are the mechanisms which determine these limitations? Is the capacity for immediate ordered reproduction invariant, and if it is not, how is it modified by past training and by the context in which it is measured? A second group of studies focuses on the mechanism of forgetting which comes into play as soon as a stimulus has been perceived and which leads to the progressive loss of what was available in immediate memory The experimental objective is to isolate what are assumed to be the elementary processes of forgetting as they occur before more permanent changes are produced by successive reinforcements The common denominator in the third category of investigations is the concern with associative processes and serial organization in short term recall Once the materials exceed the capacity of immediate memory, what are the principles determining the selective retention of some items rather than others, and the order of their reproduction? It is in this third group of studies that the role of pre experimental habits in determining selective retention and ordered reproduction becomes a major focus of interest

The Immediate Memory Span

Historical developments—The experimental procedure which focuses most directly on the *S*'s *capacity* for the reception and immediate retention of information is the measurement of the immediate memory span which is defined by the number of discrete units that can be reproduced in correct serial order after a single exposure Since the first studies by Jacobs (1887), research on the memory span has had a continuous history For a long time the memory span was treated primarily as a test of mental ability The dependence of the span on such variables as sensory modality and type of unit has been studied extensively (e g, Brener, 1940, Crannell & Parrish, 1957, Gates, 1916, Lumley & Calhoun, 1934) The conception of the memory span as a basic mental ability—it is, of course, a standard item in intelligence tests—was reflected in studies of age trends (e g Beebe, 1944, Calhoun, 1935a, 1935b, Metraux, 1944), improvement by long term practice (Gates & Taylor, 1925, Martin & Fernberger, 1929), and consistency of individual differences (Brener, 1940)

In recent years the emphasis has shifted to an analysis of the processes which determine the limits of the memory span, i e, to inquiries into the mechanisms which are responsible for the breakdown in the *S*s ability to retain and reproduce serially ordered material Such re examinations of the memory span were stimulated and aided by the development of the analytic methods of information theory The span came to be regarded as a standard measure of the organism's capacity to process and transmit information At the same time there has been increasing recognition that the memory

used in the evaluation of performance. A discontinuity of processes does not necessarily correspond to this distinction.

The functional unit of the memory span.—The accuracy of report after a single exposure depends both on the discriminations which the *S* makes among the units comprising the series and on the rate of forgetting for the successive units. The precise definition of the unit of recall is clearly a prerequisite to the measurement of the span. The units are usually specified in terms of such conventional categories as digits, letters, words, etc. The *S* is likely, however, to bring to the experimental situation habits of grouping, and responding differentially to, sequences of such units. Thus, to borrow a distinction which Underwood (1963) recently made in an analysis of the stimulus in serial learning, it becomes necessary to distinguish between the nominal unit and the functional unit in the determination of the memory span. A major contribution of communication theory to the understanding of the memory process is the development of analytic tools which permit a systematic approach to the identification and manipulation of the functional unit in immediate memory. Interestingly enough, it was not the goodness of fit of the theoretical model but rather the organism's failure to behave like a communication channel with a fixed capacity for the transmission of information which constituted the finding of greatest psychological significance.

If the memory span is limited by the amount of information that can be transmitted, the length of the span should be inversely related to the amount of information per item. As Miller (1956) has clearly shown, this is not the case: the length of the span is limited by the number of items and essentially independent of the amount of information per item. For example, in the experiment of Hayes (cited by Miller) there was only a slight decline in the length of the memory span when the amount of information per item was varied from less than 2 bits (for binary digits) to 10 bits (for a vocabulary of 1000 words). The results for the memory span are a special case of a general relationship which holds over a wide range of lengths of lists and which has been stated as follows: "The percentage of information presented that is assimilated is independent of the number of possible alternatives per message unit and is simply a function of the length of the message" (Pollack, 1953, p. 430).

The fact that the span is limited by the number rather than type of items led Miller to the suggestion that the functional unit in immediate memory is a "chunk" of information. There is a constant number of chunks which can be reproduced after a single exposure regardless of the amount of information per chunk. It follows that one can increase the amount of information in the span by grouping small chunks into larger ones with a higher informational content. The process of increasing the size of the functional unit in the span is designated as recoding. The effectiveness of recoding is

illustrated by the dramatic gain in the memory span for binary digits which is achieved when new names are attached to all possible sequences of a given length, i.e., when a translation is made from a base-two arithmetic to one with a higher base (Smith, cited by Miller, 1956).

In the example just given, recoding consisted of the attachment of differential responses or labels to sequences of discrete units. The recall of the labels mediated the reconstruction of the sequences to which they were attached. Thus the memory span may serve as a test of transfer: the size of the "chunks" reflects the distinctive differential responses to groups of discrete units established prior to the test. In order to be effective in producing an increase in the memory span the associations between stimulus sequences and differential responses or labels must be symmetrical: the successive groups of discrete units must reliably evoke the appropriate labels, and recall of the labels must in turn mediate errorless reproduction of the sequences.

Running memory span.—Verbal messages do not normally reach us in discrete series falling within the span of immediate memory but instead are likely to arrive in sequences far in excess of the span. After exposure to such a sequence some limited series of items can be reproduced consecutively but the entire sequence cannot, i.e., *parts* of the sequence will fall within the span. As the sequence continues, the particular series which can be reproduced correctly will change, i.e., some old items will drop out and some new ones will be added. These considerations lead to an important extension of the concept of memory span and of the operations by which it can be measured. In this broadened definition (Waugh, 1960) the memory span refers to the length of a subseries within a long series which can be recalled correctly after a single presentation. This subseries is measured from a reference point within the longer sequence. If the beginning of the sequence serves as the point of reference, the initial span is measured; if ordered recall is scored with the end of the series as the point of reference, the terminal span is measured. Waugh (1960) has suggested that each item in a sequence has two independent probabilities of ordered recall, and that these probabilities are determined by the distance of the item from the beginning and the end of the sequence, respectively. The two probabilities combine to the extent that the initial and terminal spans overlap. The shorter the sequence, the greater is the amount of overlap between the initial and the terminal spans and the more likely it becomes that the entire sequence can be reproduced without error. When this is the case, the measured initial and terminal spans coincide. Thus, the conventional memory span represents the special case of maximal overlap between the two span functions.

The terminal span becomes critical when a continuous sequence is in-

terrupted and the *S* is required to recall as many of the most recent items as he can. Two such situations may be distinguished (*a*) The *S* does not know the length of the sequence and hence cannot anticipate the point at which recall of the terminal items will be required The measure of recall obtained under these conditions is designated as the running memory span (*b*) The length of the sequence and hence the point at which the test will occur are known to the *S* These conditions define the known length memory span

A systematic comparison of the two types of terminal span has been presented by Pollack, Johnson, and Knaff (1959) The stimulus series consisted of randomly selected decimal digits Both the length of the sequence preceding the test and the rates of presentation were varied The running memory span was consistently lower than the known-length span, and this difference increased with the length of the series The advantage of the known-length span was maintained over an extended period of practice While both spans varied inversely with the rate of presentation, the effect of rate was much more pronounced on the known-length span than on the running span Pollack, Johnson, and Knaff attribute the superiority of the known-length span to the "behavioral strategies" which are open to the *S* under this condition When the length of the series is known to the *S*, he can reach a decision about the number of items which he will attempt to reproduce and essentially ignore the earlier part of the sequence The gain from such a strategy increases with the length of the series since considerable interference from earlier items can be avoided It is reasonable that such a strategy can be carried out more effectively at a slow than at a fast rate of presentation

The major variable determining the terminal memory span is the length of the sequence preceding recall When this variable is manipulated, the operations parallel those used to measure proactive inhibition as a function of the amount of prior learning The relationship between spans obtained without a pre recall series and with pre recall series (Waugh, 1960) is analogous to that between control and experimental measures in conventional studies of proactive inhibition Thus, the terminal span provides a method for assessing rapidly accumulating proactive interferences However, an important difference remains between the conventional experiment on proactive inhibition and the span procedure In the former the interfering items and the items to be recalled are clearly differentiated as successive lists each of which is learned to a predetermined degree, in the latter both are part of a continuous sequence and the amount to be learned is left ambiguous Hence, the behavioral strategies to which Pollack et al refer become so critical, they help to determine the effective amount of proactive interference falling on the terminal items

Results obtained in comparisons of interserial interference under conditions of intentional and incidental learning are relevant to the interpretation of the proactive influences on the terminal span. Material learned incidentally is a less effective source of interference than material learned intentionally (Postman & Adams, 1956a). This result can, of course, be attributed to the difference in the degree of learning under the two conditions. The behavioral strategy of the *S* in the terminal span experiment, who decides to 'ignore' the early part of a sequence, creates conditions similar to those which obtain in incidental learning. The individual is exposed to the items but differential responses and rehearsal are minimized.

The broadened definition of the memory span as a subseries within a sequence, with the conventional span as a special case, emphasizes the continuity between these measures and other measures of retention. When the contextual determination of the span is fully explored, including conditions of proactive facilitation as well as interference, the limits of the span are likely to vary over a wide range. The finding that both the known-length span and the running span show substantial improvement with practice (Pollack, Johnson, & Knaff, 1959) also supports the conclusion that the capacity for immediate ordered recall is far from invariant.

Differential Short-Term Forgetting

Just as the memory span has been viewed as a measure of the capacity for immediate retention, so the rate of forgetting for materials falling within the span has been used extensively in tests of hypotheses about the basic mechanism of forgetting. These hypotheses have centered around the concept of a memory trace. It is assumed that the perception of each item leaves a distinct neural aftereffect or trace. This trace is said to fade or decay gradually unless it is restored by repetition of the stimulus or by rehearsal. At any given moment in time, the probability of correct recall *depends on the degree to which the* decay of the trace has progressed. Since the trace is assumed to be short lived, the temporal course of its decay must be investigated over very brief retention intervals after a single exposure of the learning material. Once the trace has been strengthened through repetition or rehearsal, additional neural mechanisms may come into play to determine subsequent retention. For example, Hebb (1949) has proposed a dual trace mechanism—a transient reverberatory trace (an activity trace) and a more permanent trace involving structural growth (a structural trace). Thus the assumption of an unstable trace representing the aftereffects of stimulation usually entails a distinction between short term and long term memory (cf. Broadbent, 1958).

It is now apparent why materials falling within the span of immediate memory are typically used in experimental tests of the trace hypothesis

First, it seemed plausible that the effects of a single exposure to a series of discrete units such as digits would be limited to a transitory activity trace (Hebb, 1949, p 62) It is important to note, however, that in a recent study Hebb was led to the conclusion that "a single repetition of a set of digits . . produces a structural trace which can be cumulative" (1961, p 43) In Hebb's experiment, repetitions of the same series of digits were interspersed in a sequence of tests of the *S*s' digit span The recall of the recurrent series improved steadily as a function of the number of repetitions, both with and without correction for errors In the light of Hebb's findings it would no longer be reasonable to equate the effects of a single repetition of a series with a pure activity trace

Other reasons favoring the use of memory span materials in tests of the trace hypothesis are methodological For purposes of such tests, recall immediately after a single presentation should be high, so that there is an opportunity to observe gradual changes in the probability of recall during a short retention interval This condition is, of course, met when the length of the learning series does not exceed the immediate memory span In actual practice the span is not normally determined in studies of short term retention but series of fixed length are used which are known to be within the span or at most to exceed it slightly Finally, memory span materials have the advantage of being composed of homogeneous units When a series is recalled, different amounts of time have elapsed since the presentation of each of the items constituting the series Accordingly, the period of trace decay is a function of the position of an item in the series The same holds true for the position of an item in recall Comparisons of the rates of forgetting for individual items or groups of items within a series are greatly facilitated when the units are of homogeneous difficulty We turn now to the experimental procedures which have been devised to test the hypothesis derived from trace theory that forgetting is a direct function of the time interval between presentation and recall

Rate of presentation and recall—The experimental isolation of time as an independent variable offers major difficulties It is necessary to vary the length of the retention interval while holding interpolated activities constant and minimizing rehearsal One procedure designed to meet these requirements is the manipulation of the rates of presentation and recall The slower these rates the longer is the average time interval between the presentation and reproduction of an item According to decay theory, amount retained should vary directly with the rates of presentation and recall The relationship predicted by decay theory has been found in a series of experiments on the retention of series of digits presented and recalled at different rates (Conrad, 1957, Conrad & Hille, 1958, Fraser, 1958) The faster the rate, the higher was the proportion of series recalled

correctly Since speed of learning in conventional anticipation experiments varies inversely with the rate of presentation, the results were interpreted as strong support for the decay hypothesis It should be noted, however, that in other studies using different experimental procedures the expected inverse relationship between rate of presentation and memory for digits has been found (Moray, 1960, Pollack, Johnson, & Knaff, 1959), or the span was found to be independent of the rate of presentation (Peatman & Locke, 1934) These results are not necessarily contradictory since the rates of recall were not controlled in the latter studies The findings of Conrad and Hille (1958) show that significant variations in the amount of recall are obtained only when the differences in rate apply to both presentation and reproduction

There is also some evidence that the effects of rate may vary with the type of material In an early investigation by Bergstrom (1907) recall of verbal items was an inverse function of the rate of presentation but the relationship was much more pronounced for words than for letters Presumably reductions in the rate of presentation provided an opportunity for the formation of interitem associations which were more effective for words than for letters On the other hand, Fraisse (1937, 1945) found a direct relationship between rate of presentation and accuracy of reproduction for sound patterns but no effect for series of consonants According to Fraisse, a rapid rate of presentation favors the perception of patterns which is critical for the reproduction of series composed of identical elements such as sounds, perceptual grouping loses its importance when each member of the series is a distinctive unit (e g , a consonant or digit) It is reasonable to conclude that the available evidence on the relationship between rate of presentation and recall is too fragmentary and inconsistent to have a decisive bearing on the evaluation of the decay hypothesis

The procedure of varying the retention intervals for otherwise comparable series of items by manipulation of the rates of presentation and recall is not free of methodological difficulties It is true that the critical time intervals are varied with the amount of interpolated learning held constant The opportunities for rehearsal, however, clearly remain uncontrolled Conrad and Hille (1958) suggest that to the extent that rehearsal occurs, it should work against rather than in favor of their hypothesis This point is valid only if one assumes that under the conditions of their experiment (*a*) rehearsal did not interfere with the perception of successive units, and (*b*) predominantly correct responses were rehearsed It is known, however, that the opportunity for rehearsal does not necessarily improve retention since errors as well as correct responses may be practiced (Postman & Phillips 1961 Rohrer 1949) This is quite likely to be true for successive series of digits since errors of transposition and substitution frequently occur

in the reproduction of such materials Thus, length of the average retention intervals was not the only variable in these experiments, and the results cannot be considered conclusive

Order of recall—Manipulation of the order of recall provides another method of varying the retention interval for individual items within a series The assumption is that in the reproduction of a series the items which are recalled first are also recalled best because the period of delay between presentation and test is shortest for them In ordered recall the total period of delay is determined both by the position of an item in the series and by its position in recall The effect of recall order per se can, however, be evaluated when stimulation is simultaneous and recall is successive This is the case in the split-span technique used by Broadbent in which pairs of digits are presented simultaneously to the two ears, one arriving at the left ear and the other at the right ear Reproduction of both sets of digits is possible but *S*s characteristically report all the digits presented to one ear before any of those presented to the other ear, and those reported first are recalled best Reproduction in the order of arrival becomes possible only if the rate of presentation is relatively slow (Broadbent, 1954)

In the theoretical analysis of these results, Broadbent (1957b, 1958) has distinguished between two mechanisms which function in the reception and reproduction of information presented to the organism the *S* system which represents a stage of neural activity at which information can pass simultaneously, and a later stage designated as the *P* system which can pass information only successively The *S* system may be regarded as a mechanism of short term storage whereas the *P* system refers to ongoing perceptual activity When information is delivered through more than one channel, the organism's ability to shift from one channel to the next is limited Items which cannot be attended to at the time of their arrival remain in the short term store until a shift of attention becomes possible Such shifts in attention are favored by a slow rate of presentation (Broadbent, 1954), by a staggered rather than strictly simultaneous delivery of the two messages (Moray, 1960), the perceptual similarity of the concurrent series (Broadbent, 1956, Broadbent & Gregory, 1961), and the existence of pre experimental associations among items in the simultaneous messages (Gray & Wedderburn, 1960) As Peterson (1963) has pointed out, it is not always possible to determine whether the groupings evident in the *S*'s report occurred at the time of stimulation or at the time of recall In any event, the fact that with simultaneous stimulation the order of recall frequently does not correspond to the order of arrival at the sense organs has been used to assess the efficiency of the short term store in which "unattended" items must remain before they are recalled

There are three essential findings (*a*) In general, the items recalled first

are recalled best, as long as the *S* is free to reproduce the items in any order he wishes. (*b*) When the possibilities of perceptual confusion are minimized, as for example by simultaneous presentation to eye and ear, retention after simultaneous stimulation is at least as good as after successive stimulation through a single sensory channel (Broadbent, 1956). When successive responses are made to simultaneous stimuli, the rate of forgetting is no faster than after successive presentation. (*c*) Short term storage is possible, however, for only a strictly limited period of time. If the response to the second channel is delayed by a second or so, retention is appreciably reduced (Broadbent, 1957a). This is interpreted to mean that the stimulus trace fades rapidly unless it is restored by rehearsal, or as Broadbent (1958) prefers to put it, unless there is circulation from the *S* system through the *P* system and back again. When the *P* system is ' kept occupied' responding to another channel, extremely rapid forgetting is in fact observed.

When two short series are presented successively, the order of recall can be manipulated by instructions. Again, items recalled first are usually recalled best. However, the relationship between the order of presentation and the order of recall as well as the *S*'s expectations about the required order of reproduction now become significant variables. When the order of presentation and the order of recall are reversed, the series recalled first no longer has an advantage, i e , the effects of the order of presentation and the order of recall appear to cancel each other (Kay & Poulton, 1951). Instructions about order of recall given prior to presentation produce a higher level of performance and a different distribution of errors than instructions after presentation (J. Brown, 1954, Kay & Poulton, 1951). It is clear that the effects of the order of recall are modified by interitem interferences and by the *S*'s set during exposure and at the time of the test.

When accuracy of reproduction does vary with order of recall, it is hazardous to conclude that this relationship reflects differences in the time of decay for the traces of individual items. It is true that the later an item is recalled, the longer is the time interval which has elapsed since original exposure. However, it is equally true that the later an item is recalled, the greater the amount of recall activity which has been interpolated between its original presentation and its reproduction. There is evidence that interpolated recall of a set of items interferes with the retention of material which still remains to be reproduced (J. Brown, 1954) and also with the acquisition of new materials (Poulton, 1953). Such interferences cannot readily be distinguished from the variations in the length of the retention interval correlated with order of recall.

Similar difficulties arise in the interpretation of results obtained by the

method of *poststimulus cuing* When this method is used a series of items is exposed but the *S* is required to reproduce only a designated part of the material Differential cues such as lights or sounds are used to indicate the part of the material to be included on a given test For example, in the experiment of Anderson (1960), lists of 12 digits were presented in groups of 4, and on different trials *S*s were required to reproduce either 4, 8, or 12 items For each segment of the series, the percentage correctly reproduced was inversely related to the length of the recall task The detrimental effect of lengthening the task may be a function of interpolated recall as well as the increase in the average length of the retention interval

One consequence of the relationship between the order and accuracy of recall is that the total recall score systematically underestimates the amount of information available at the beginning of a test of retention Since accuracy declines as recall proceeds, not all items initially available will be correctly reproduced It is possible, however, to use partial reports obtained by means of poststimulus cuing to estimate the total amount of information available for recall after a given retention interval This method of analysis was used in a study by Sperling (1960) Groups of letters were exposed in various spatial arrangements, such as three rows of four letters each The *S*'s reports were limited to a part of the array at a time, with poststimulus cuing by tones of varying pitch The partial reports were considered as random samples of the information which the *S* was able to recall To estimate the total amount of information available, the number of items correct in the partial report was multiplied by the number of equally probable, nonoverlapping reports which could be required For zero delay between exposure and report, these calculations yielded estimates of available information far in excess of those yielded by conventional measures of the memory span The available information appeared to decay rapidly, however, and 1 sec after exposure dropped to the level of the conventional span In a similar study, in which visual rather than auditory poststimulus cues were used, Averbach and Coriell (1961) concluded that the duration for which material remains in the immediate memory store may be as short as 0 25 sec This estimate includes adjustments for 'nonselective read out' (something like uncontrolled rehearsal during the retention interval), as well as for the time occupied by the detection of the poststimulus cue and the 'selective read out' of the partial report called for by that cue

The distinction made in these analyses between what is available in immediate memory and what is reproducible on a given test emphasizes the fact that retention is a theoretical construct which cannot be equated with any one measure such as recall This conclusion applies, of course, to long-term as well as to short term retention The shorter the time interval, how-

ever, the more likely it becomes that the effects to be measured change significantly during the period of measurement and hence are influenced by the duration of the test.

Retroactive and Proactive Inhibition in Short-Term Retention

Experimental design.—Ever since Muller and Pilzecker (1900) proposed their perseveration hypothesis, it has been clear that the time interval immediately after exposure to a set of materials is of critical importance for the assessment of the temporal development of the hypothetical memory trace. According to the perseveration hypothesis, this is the interval during which the trace is consolidated. From the point of view of decay theory the rate of forgetting during this interval should reflect the progressive fading of the trace.

Critical tests of such hypotheses have been seriously complicated by the difficulties of controlling and evaluating the effects of the activities intervening between presentation and recall. There are two difficulties. First, rehearsal may reduce or prevent forgetting. As was pointed out above, however, the opportunity for rehearsal does not necessarily improve retention. In any event, in tests of the decay hypothesis it is essential to prevent rehearsal during the retention interval. Second, when an activity other than the rehearsal of correct responses fills the interval, conditions of interference may be created which mask the normal development of the trace. According to decay theory such interferences should, however, be independent of the degree of similarity between the original task and the interpolated activity. This assumption does not necessarily hold for all types of trace theory since one may choose to posit an interaction among traces (Koffka, 1935; Kohler & von Restorff, 1935).

In order to achieve control over the conditions of interpolation without sacrificing the advantages inherent in the use of a single exposure, Robinson (1927) developed a modification of the conventional memory-span method. A series of consonants which was somewhat in excess of the average memory span was presented. For purposes of analysis the first and second half of the series were treated as original and interpolated learning, respectively, and the similarity relationships between these sublists were systematically manipulated. Defining similarity in terms of the number of identical elements, Robinson found that recall for the first half of the span varied directly with the number of identical elements in the two halves. The U-shaped relationship with similarity expected on the basis of the Skaggs-Robinson hypothesis was not obtained nor did it occur reliably in subsequent studies in which the range of similarity was extended by introduction of varying numbers of digits into the second half of the list (Harden, 1929; Kennelly, 1941). The effects of similarity per se remained uncertain,

however, because (*a*) the use of identical elements in the two halves of the span provides an opportunity for rehearsal of the critical items, and (*b*) variations in the proportion of consonants and digits leave uncontrolled the difficulty of the "interpolated" series and of the total span (cf. Young & Supa, 1941).

The renewal of interest in decay theory has led to the development of experimental procedures which follow Robinson's method of a single continuous exposure of the original and interpolated series but which have been modified so as to permit tests of specific hypotheses about the properties of the memory trace. When the decay of the trace is under study, the interpolated activity is introduced in order to control stimuli to which the *S* is exposed during the retention interval and to minimize rehearsal. For these purposes it is not necessary, and indeed undesirable, that the interpolated activity consist of a learning task. Interpolated learning may produce errors of generalization which are to be separated sharply from forgetting attributable to decay of the trace. Negative transfer between successive tasks must, therefore, be avoided unless the factor of response confusion is explicitly under investigation. In the light of these considerations the first and second half of the series are clearly differentiated and the *S* is instructed to learn and recall only the former. However, incidental learning of the interpolated materials obviously cannot be prevented.

Conditions of interference.—A study by Brown (1958) will serve as a reference experiment exemplifying this type of design. The purpose of this experiment was to test a deduction from trace theory concerning the relationship between the length of the learning series and rate of forgetting. The larger the number of items in a series, the longer is the average time of delay between the presentation of an item and the end of the series. Since the trace is assumed to decay progressively during this period of delay, the strength of the average item at the end of exposure should vary inversely with the length of the series. It follows that amount retained after a constant retention interval should also vary inversely with the length of the series, provided rehearsal is prevented. In Brown's experiment the first part of the stimulus series, which was to be recalled by the *S*, consisted of pairs of consonants. The length of this series was varied between one and four pairs. For an experimental group the 5 sec. retention interval was filled with the recitation of five pairs of numbers, whereas a control group had no interpolated activity and was presumably free to rehearse.

As expected, the amount of retroactive inhibition varied directly with the length of the series. In further experiments using the same kind of experimental arrangement Brown found that with a 5 sec. retention interval the degree of similarity between the learning items and the interpolated stimuli did not influence the amount of retroactive inhibition, nor was proactive

inhibition a function of interserial similarity In an earlier study Pillsbury and Sylvester (1940) bad also reported substantial amounts of retroactive and proactive inhibition with a 10-sec retention interval and dissimilar tasks While the results for retroactive inhibition are consistent with the theory, the occurrence of proactive inhibition cannot be predicted on the basis of a decay bypothesis Hence, Brown attributes proactive effects entirely to confusion among responses at recall Retroactive inhibition, which was considerably greater than proactive inhibition, presumably reflects both the decay of the trace and reproductive inhibition at recall This is essentially a two-factor theory (Melton & Irwin, 1940, Melton & von Lackum, 1941), with a decay factor substituted for unlearning It is interesting to note that it receives apparent support from the greater magnitude of retroactive than proactive effects just as does the classical two-factor theory

It is a fact that an interpolated activity filling a very brief retention interval (of the order of 5 to 10 sec) produces substantial amounts of retroactive inhibition According to decay theory the interpolated activity reduces retention because it prevents restoration of the fading trace by rehearsal If this analysis is correct, it follows further that the detrimental effects should vary directly with the length of the interpolated activity, i e , with the extent of reduction in rehearsal time This implication was tested in a recent study by Conrad (1960a) whose experimental procedure represents an ingenious modification of the retroaction paradigm for the purpose of evaluating the influence of minimal amounts of interpolation The materials to be reproduced were series of digits In the experimental conditions the *S*s were instructed to precede their reproductions with a single additional digit, viz , "zero" This single interpolated response which occupied only a fraction of a second produced an amount of retroactive inhibition comparable to that obtained with an interpolated task filling the entire retention interval in Brown's experiment Even more important is the finding that the interpolated response results in equal amounts of interference whether it occurs at the beginning or the end of a retention interval during which rehearsal is possible In fact, the retroactive effect occurs when *S*s are instructed to rehearse the learning series prior to giving the interpolated response Contrary to the assumptions of decay theory it is not the length of the interval but the fact of interpolation per se which is responsible for the retention decrement in this situation

Analysis of errors —A decay theory predicts that responses become increasingly less available as a function of time and does not specify the conditions under which overt intralist and interlist intrusions will occur Additional principles are required to account for the occurrence of such errors and the question arises of whether such principles can be developed

within the framework of decay theory or lead to a multiple process conception of forgetting

A first problem is posed by *misplaced responses, i e*, correct items reproduced in an incorrect serial position Such errors are quite common in tests of immediate memory, in reproductions of eight digit series, for example, transposition errors account for about 50% of all overt errors (Conrad, 1959) The fact that the correct item itself remains available implies that its trace had not faded altogether, yet the serial position has been forgotten Consequently, a distinction is introduced between "order elements" and "item elements" in the stimulus trace (Conrad, 1959), with the former assumed to have less temporal stability than the latter In a similar vein Brown (1958) has suggested that there is usually a higher degree of redundancy in the characteristics of individual items than in the specification of the serial order After partial decay of the memory trace correct reproduction of individual items is, therefore, still possible whereas the serial order is lost Brown extends this analysis to account for the fact that the memory span appears to be limited by the sheer number of items and to be independent of the amount of information per item (Miller, 1956) Since the span is defined by the length of the sequence which can be reproduced correctly, it may be the informational content of the order rather than of the individual items which is primarily responsible for the invariant length of the memory span

A second systematic issue arises in the interpretation of interlist intrusions when, as is customary, several tests of immediate memory are given in succession It is a fact that under these conditions there are frequent serial order *intrusions, i e*, intrusions of items from corresponding serial positions in other lists (Conrad, 1959, 1960b) The amount of forgetting attributable to the decay of memory traces becomes ambiguous if at least some of the retention loss is produced by the intrusion of competing responses from other lists An alternative interpretation of the facts is, however, possible As Conrad points out, it is not necessarily the case that interlist intrusions lead to forgetting It is possible that forgetting of the correct response in a given serial position is a condition for the occurrence of serial order intrusions Having forgotten the correct response, the *S* "selects" the intruding item, which had been recalled and rehearsed, because it is appropriate for a given serial position

To decide between these alternatives, Conrad (1960b) varied the time interval between successive series A range of intervals between 15 and 40 sec was used The critical question was whether changes in the frequency of intrusions as a function of time interval would be reflected in proportionate changes in the amount of forgetting Conrad found that the number of intrusions declined steadily as a function of the interlist interval

whereas the amount recalled remained invariant He concluded that the occurrence of interlist intrusions was a consequence rather than a condition of forgetting The lack of correlation between amount of recall and frequency of interlist intrusions is, of course, a well established fact in studies of interserial interference (e g , Melton & Irwin, 1940, Underwood, 1945, 1949) In these studies the independent variables have been the degree of learning for successive lists and the length of the retention interval Regardless of whether Conrad s interpretation is accepted, his results are of interest in showing that the lack of correlation between amount of forgetting and frequency of intrusions also obtains when the time interval between successive lists is varied

The experimental designs discussed in this section do not appear to have produced conclusive evidence for the decay theory of short term memory The methodological difficulties mentioned earlier continue to await solution No satisfactory procedure has been developed to separate the effects of time interval per se from those of uncontrolled rehearsal on the one hand and interfering interpolated tasks on the other As long as such procedures cannot be devised, the short term forgetting functions cannot be coordinated with the decay of the hypothetical memory trace In addition, the fact of short term proactive inhibition poses a serious explanatory problem to trace theory

The experimental results leave little doubt about the occurrence of retroactive and proactive effects in the recall of short series of discrete units after very brief retention intervals Direct comparisons with the main body of findings on interserial interference are often difficult, especially because of the absence of adequate measures of the degrees of learning There is nothing in the results, however, to compel the conclusion that basically different principles govern short term and long term interference

Short Term Retention During a Continuous Task

In all the experimental procedures discussed thus far, the presentation of the materials and the test of recall are clearly separated in time While this arrangement is optimal for purposes of experimental control, it is clear that it is not representative of many practical situations in which short term memory must function When the organism is required to respond to a continuously changing environmental situation, exposure to new items and recall of old ones are likely to occur in various patterns of alternation In some cases newly received information may have to be recalled and used immediately after presentation in other cases reception and recall of other items intervene Thus fluctuating amounts of retroactive and proactive inhibition fall on the recall of successive items Such situations provide information about what has been called the storage capacity' of short

term memory, i e, the average amount that can be retained while old items are being discarded and new ones are being added The distinctive feature of the experiments to be discussed in this section is that the limits of short term retention are tested by varying the storage loads imposed on the *S*'s memory during a continuous task Such a task may simply involve recall or recognition of items presented previously or performance of a series of responses guided by the retention of preceding signals We shall refer to these two types of situation as continuous memory tasks and continuous instrumental tasks

Continuous memory tasks—The first reference experiment to be considered is that of Lloyd, Reid, and Feallock (1960) The learning materials were words, and recall was tested by presenting as a cue the name of the class to which the word belong (e g, "tree" served as a cue for the recall of "pine") Presentation and recall of items occurred in various serial orders during an experimental sequence, i e, presentation and recall were separated by varying numbers of other presentations and recalls The major experimental variable was the "average storage load," which was defined by the average number of items which the *S* had to remember in order to be prepared for the tests as he progressed through the sequence A direct relationship between the average storage load and the number of errors was found The same relationship holds when unrelated words are used as stimuli and letters paired with the words are the cues to recall (Lloyd, 1961) The presence of significant practice effects as a function of experimental days suggests that *S*s 'learn how to remember' in this type of complex situation Other factors such as number of item classes or number of different items per class had little or no influence on performance, and the investigators concluded that the average load was the major variable determining recall

In a subsequent study (Reid, Lloyd, Brackett, & Hawkins, 1961) a second major variable determining accuracy of recall during the continuous task was identified, viz, the average load reduction Amount of load reduction refers to the number of items which the *S* is required to recall at each test point With average storage load (number of items to be remembered) held constant, errors decreased as a function of the average amount of load reduction per recall point The improvement in performance may be the result of (*a*) the reduction in the number of tests which serves to decrease the amount of interpolated activity between presentations and recalls, and (*b*) the opportunity for associative clustering to aid recall when several items from the same class are called for An experimental decision between these alternatives can be sought As the authors point out, the method permits the systematic manipulation of selected experimental variables with the contextual conditions held constant

A related procedure for measuring performance in continuously changing memory tasks has been developed by Yntema and Mueser (1960, 1962) A series of messages is presented to the *S*, each of which indicates that one of a number of variables (attributes of objects) is in a particular state and will remain in that state until a change is announced in a subsequent message The messages are interrupted at random intervals, and the *S* is required to answer a question about the current state of one of the variables "States" refer to members of classes such as shapes or directions, or members of aggregates such as consonants Again storage load is a significant determinant of the accuracy of recall, i e , the proportion of the questions answered correctly decreases as a function of the number of variables to be remembered Performance also depends on the susceptibility of successive messages to associative interference Thus, the error rate is higher when the messages all refer to the same attribute of different objects (e g , all specify the shape of one of a series of objects) than when they refer to different attributes of the same object As the authors point out, the former arrangement conforms to the A-B, A-Br paradigm in studies of interserial transfer in which identical stimuli and responses are re paired in successive lists This paradigm is known to maximize negative transfer as compared with experimental arrangements in which either stimuli or responses are different in successive tasks (Postman, 1962c, Twedt & Underwood, 1959) It appears that principles of interserial transfer can be used to predict performance in continuous memory tasks This continuity of principles is also apparent in other characteristics of performance the probability of a correct response declined with the number of items intervening between the message and the test of recall, and many of the overt errors were analogous to interlist intrusions, i e , *S*s tended to substitute a previously correct state for the one announced in the most recent message

A technique for measuring recognition in a continuous task was recently reported by Shepard and Teghtsoonian (1961) The learning materials were long series of three digit numbers, with each number occurring twice in the series The *S*'s task was to identify each number as either new or old New and old numbers followed each other in a random sequence The major independent variable was the length of delay, i.e , the amount of intervening old and new material, between the first and second presentations of a number As was to be expected, the probability of a correct response declined as a function of the number of intervening items What is noteworthy is the length of the delays after which better than chance performance remained possible even after 60 intervening items the probability of a correct recognition was higher than that of classifying the new item as old Shepard and Teghtsoonian calculate that the information

carried by the *S* was 32 bits as compared to 23 bits recalled after one exposure to a series of digits. Recognition appears to be a considerably more sensitive measure than recall of short term as well as long term retention. This fact deserves emphasis since recall has been used almost exclusively in tests of the decay theory with verbal materials. Shepard and Teghtsoonian, who favor a trace model of forgetting, point out that recognition may remain possible even after there has been some fading or "diffusion" of the trace. Trace theorists appear to face a major analytic problem in coordinating variations in the degree of decay with differences among measures of retention.

Continuous instrumental tasks.—We turn now to situations in which the degree of short term retention for successive signals determines the accuracy of performance in a complex serial task. The available data indicate that forgetting under these circumstances is considerably more rapid than in a conventional memory situation. In the experiment of Mackworth and Mackworth (1959) complex displays were presented, and the *S* had to determine by visual search the particular part of the display to which he had to respond at any point during the paced presentation. Advance cues to successive objectives were given, and the number of moves separating the cue from the appropriate response was varied systematically. Advance cues could be retained for only a very limited span of time, they were effective only if they preceded the objective by no more than one or two moves. The temporal range over which advance cues were effective varied inversely with the rate of responding and could be increased by grouped presentation. In a related experiment J. F. Mackworth (1959) compared the effectiveness of verbal and non verbal cues (labeled and unlabeled lights) at different periods of delay. Nonverbal cues were more effective at short delays whereas at the longer delays the difference was in the opposite direction. Mackworth suggests that the memory trace evolves through two successive stages—a brief perceptual trace followed by a secondary trace which has a verbal component. The accrual of the verbal component, which is required for durable retention, depends on rehearsal. At the shorter delays the nonverbal cues are more effective because the need for rehearsal is eliminated, at the longer delays such cues become ineffective because of the rapid decay of a perceptual trace which is not reinforced by a verbal component.

The extension of experimental arrangements to include retention during continuous tasks as well as for circumscribed series parallels the developments in the study of the immediate memory span. In both cases there is an attempt to evaluate the dependence of retention functions on the verbal and behavioral contexts in which they are measured. The difference in the efficiency of short term retention in continuous memory tasks and con-

tinuous instrumental tasks clearly illustrates the importance of contextual variables

Short-Term Retention of Single Items

The basic mechanisms of forgetting postulated by general theories of memory typically embody a hypothesis about the conditions governing the loss of retention for individual items This is clearly true of the type of trace theory that we have been considering as well as for other versions of trace theory such as the hypothesis of autonomous memory change (Wulf, 1922, Koffka, 1935) The specific mechanisms used in the elaboration of the interference theory of forgetting, such as unlearning and response competition, also apply in principle to the recall of individual associations There are, of course, some important hypotheses about the conditions of retention which have reference to the relationship among a series of items, e g , those concerned with the spontaneous recovery of intralist generalization (Gibson, 1940) and with the effects of interitem associations (Deese, 1959a) It would appear to be possible to divide hypotheses about the mechanism of forgetting into those which entail predictions about the retention of single units and those which in principle require the consideration of multiple units of learning Difficulties of interpretation often arise when hypotheses formulated to account for the retention of single units are tested in experiments involving the learning and retention of multiple units Thus, when a series of items is used in a study concerned with the temporal development of the stimulus trace, the analysis is complicated by the need to take account of intraserial associations and interferences These difficulties are, of course, avoided when the learning material is restricted to a single item Experimental arrangements permitting the use of single items were first developed in studies of memory for form and more recently with verbal materials

Memory for form —The hypothesis, first formulated by Wulf (1922), that the memory trace of visual forms changes progressively toward greater simplicity and symmetry applies to the retention of single perceptual units When the learning materials consist of a series of different forms, retention may be influenced by generalization among the items in the series Early in the experimental history of this problem Gibson (1929) presented evidence for "figure assimilation," i e , loss of differentiation among the forms in a series Nevertheless, most subsequent investigators continued to use series of stimulus forms, apparently in the interests of sampling a variety of perceptual antecedents of memory change Thus, intraserial generalization remained an uncontrolled source of variation Failure to take account of interitem interferences can also result in systematic biases during the testing procedure itself as became clear when the method of recognition rather than reproduction was used to measure retention.

Ever since the demonstrations of Hanawalt (1937) and Zangwill (1937) that recognition provides a more sensitive measure of memory for form than does reproduction, there has been increasing reliance on recognition tests in experimental investigations of the hypothesis of autonomous change The procedure is exemplified by the study of Hebb and Foord (1945) in which *S*s were exposed to two visual forms and later were required to choose the correct figures from among sets of alternatives which varied along theoretically relevant dimensions of change, i e, closure of a broken circle and sharpening of an angular figure While some of the recognition tests showed significant retention losses, there was no evidence for progressive changes in accordance with Gestalt principles These results can no longer be considered conclusive, however, since it has been shown that in a recognition procedure such as that used by Hebb and Foord the *S*'s choices are systematically influenced by the order in which the alternatives included in the test are presented For example, when the original figure is a circle with a gap, the direction of the error in recognition depends on the particular place at which the *S* begins the inspection of the graduated series of incorrect alternatives If the inspection begins with gaps smaller than the original, the errors of recognition are in the direction of 'closure', if the test figures seen first have gaps larger than the original figure, the errors are likely to be in the opposite direction (Carlson & Duncan, 1955) It is clear that exposure to multiple test items constitutes a significant source of interference

Interitem interference is effectively eliminated if only a single item is used both during learning and in the test of retention In studies of memory for form this requirement can be met by an adaptation of the psychophysical method of successive comparison in which the standard and comparison stimuli are identical, but the *S* is forced to make a judgment of some specified attribute To gauge progressive changes in the memory trace of the standard, the time interval between the standard and the comparison stimulus is varied *This is the procedure frequently used in the measurement* of the time order error, and its applicability to the investigation of the memory processes is apparent (Pratt, 1933, 1936)

The method of successive comparison was first applied to the problem of progressive memory change by Irwin and his associates (Irwin & Seidenfeld, 1937, Irwin & Rovner, 1937) However, some more recent studies (Crumbaugh, 1954, Karlin & Brennan, 1957) in which the possibilities of control afforded by this method were fully exploited will serve as more convenient reference experiments In these experiments each *S* was exposed to only one figure which served as both standard and comparison stimulus The time interval between the standard and comparison stimulus was varied systematically The use of a single item made it possible to investigate changes over very short intervals of time, the intervals ranged between 0 3

and 12 sec in the experiment of Crumbaugh, and between 1 and 8 sec in that of Karlin and Brennan Repeated observations were made at each of the intervals The results of the experiments will not be reviewed here although it is possible to say that neither of them provides clear support for the hypothesis of autonomous change [For a comprehensive review of the experimental literature on memory for form, see Riley (1962)] Instead, we wish to emphasize the methodological points which are illustrated by these studies

These methodological points are the following (1) The use of a single item eliminates intraserial associations and interferences, permitting the test of hypotheses which in principle apply to the retention of single units The hypothesis of autonomous change in the memory trace is a case in point (2) When recognition is the criterion of retention, the operations for measuring temporal trends in retention are continuous with those of the psychophysical method of successive comparison Other psychophysical procedures may be similarly adapted for quantifying systematic changes in the retention of selected characteristics of a stimulus For example, Lovibond (1958) used the method of average error in testing for progressive changes toward symmetry in the long term retention of an asymmetrical figure (He found no evidence for such changes) (3) When a psychophysical procedure such as the method of successive comparison is used in the measurement of short term retention, the activity filling the interval between the standard and the comparison stimulus should be added to the variables which must be brought under experimental control In the studies cited above, no such control was attempted It is clear that verbal associations to the stimulus figure can be rehearsed during the interval between presentations Since verbal mediators are frequently invoked to account for whatever systematic error tendencies are observed (Riley, 1962), control over interpolated activities becomes essential It is to be noted that when additional stimuli are interpolated between the standard and the comparison stimulus, the operations are continuous with those for the measurement of retroactive inhibition (Postman & Page, 1947, Pratt 1936) Thus, the adaptation of conventional psychophysical procedures makes it possible to measure retroactive inhibition in the recognition of single items

Recall of single verbal items —While it would be possible in principle to use the method of successive comparison to measure the temporal course of recognition of verbal units such a procedure has not been used Recall after varying intervals of controlled activity designed to prevent rehearsal has been used instead The experimental paradigm has, therefore, been essentially that for the measurement of short term retroactive inhibition, with maximal dissimilarity between original and interpolated activities These studies are discussed in the present section rather than in the pre-

ceding one reviewing studies of retroactive inhibition in order to place major emphasis on the use of single verbal units This classification is, of course, arbitrary and made only for convenience of exposition

The reference experiments are those of Peterson and his associates The basic procedures were developed in the study by Peterson and Peterson (1959) Recall of single consonant syllables of low association value was studied Immediately after presentation of the syllable the *S* began counting backward and continued to do so until the signal for reproduction of the items was given Six different retention intervals ranging between 3 and 18 sec were used Each *S* was tested eight times at each interval Retention decreased rapidly as a function of the retention interval, declining to less than 10% after 18 sec The question arises at once whether the extremely rapid rate of forgetting is attributable to the accumulation of massive amounts of proactive inhibition during the repeated tests with the same *Ss* Peterson (1963) has been inclined to discount this possibility In his experiments successive blocks of six tests each showed some increase in recall as a function of practice at the shorter retention intervals but no evidence for cumulative proactive inhibition Such an analysis is not conclusive, however, since degree of learning for successive items increases as a result of practice, and these positive effects may be expected to counteract the cumulative effects of interference The question has since been clarified in an experiment by Keppel and Underwood (1962) which was designed explicitly to investigate proactive inhibition in the retention of single items The results of this experiment show that proactive inhibition does, indeed, develop progressively as a function of the number of prior items learned in a situation such as that used by Peterson and Peterson To demonstrate this effect unequivocally it is necessary to chart the course of proactive inhibition from the beginning of the experimental session and to use degrees of learning which permit sufficient variation in the amount of subsequent retention In agreement with the trends in long term memory (Postman, 1962b) the amount of proactive inhibition increases with the length of the retention interval (between 3 and 18 sec), and the number of prior items interacts with the length of the interval in determining the amount of interference Thus proactive inhibition appears to be governed by the same principles in short term and long-term retention [2]

There is good reason to believe, then, that the same principles of inter-

[2] The relationship between the number of prior items and the amount of interference may vary with the materials and the method of measuring retention In an experiment by Murdock (1961) the learning materials were monosyllabic English words and the *S* was required to recall only the last member of a given series instead of being tested on each successive item Under these conditions a U shaped relationship between the number of prior items and recall was obtained

ference apply to the short term retention of single items and to the long-term retention of lists The available evidence suggests that other determinants of retention likewise have parallel effects in the two types of situation (1) When rehearsal is permitted, the level of the retention curve is a function of the number of repetitions (Peterson & Peterson, 1959) The probability of recall varies directly with the degree of overlearning (2) When an item is presented twice, recall increases with the spacing interval between the two presentations (Peterson & Peterson, 1961) (3) Retention increases as a function of the meaningfulness of the item Words are recalled better than nonsense syllables, and the retention of nonsense syllables in turn varies with the association value (Peterson, Peterson, & Miller, 1961) A recent study by Murdock (1961) examines the role of pre experimental response integration in the short term retention of single units Using the same procedure as Peterson and Peterson, Murdock found considerably higher recall for single monosyllabic words than for trigrams However, when clusters of three monosyllabic words and single trigrams were compared, the retention curves for the two types of materials overlapped closely Murdock suggests that the rate of forgetting is invariant for a given number of "chunks" The differences in retention may, however, be a function of variations in degree of learning (cf Underwood, in press)

In a theoretical analysis of the conditions determining the recall of single items Peterson (1963) has proposed a distinction between two types of associative processes which combine to determine the retention of an item over a short interval of time, viz, 'background conditioning" and "cue learning" The former refers essentially to the association between the response and the general context of the experimental situation, whereas the latter denotes the acquisition of the particular sequences of letters prescribed by the experimenter The index of background conditioning used by Peterson is the recall of the first letter of the item, whereas cue learning is measured by the degree of serial dependency between successive letters, i e , the mean dependent probabilities of the second and third letter being recalled, given correct recall of the preceding letter These two measures do not necessarily vary together as a function of the experimental conditions Specifically, the dependent probabilities increase with meaningfulness whereas recall of the first letter does not On the other hand, spacing of repetitions with items of low meaningfulness favors recall of the first letter but does not influence the dependent probabilities Meaningfulness favors cue learning whereas spacing determines the proportion of background elements to which the response can be conditioned (cf Estes, 1955)

The fact that meaningfulness and spacing have differential effects on the two measures of performance does not in itself seem to require the assumption of two distinct associative processes The assumption that only the first letter is conditioned to the background and in turn serves as a cue to the recall of other letters is ad hoc and becomes quite implausible when highly integrated responses are learned Since meaningfulness and response integration are closely related (Underwood & Schulz, 1960), the high degree of serial dependency in the recall of meaningful items may be attributed to the transfer effects of preexperimental habits rather than superior cue learning (see Murdock's findings mentioned above) The fact that, with meaningfulness held constant at a low value, spacing influences the recall of first letters but not the degree of serial dependency does not necessarily point to a process of background conditioning as distinct from response integration If the integration of items of low meaningfulness proceeds gradually and in a sequential order, it is reasonable that the recall of first letters should be more sensitive to the small beneficial effects of distribution than the dependent probabilities. In short, we find nothing in the available data on the recall of single items which calls for a separation between background conditioning and cue learning This is not to deprecate the importance of contextual stimuli in the recall of verbal items However, when only a single item is learned and recalled, it is difficult to make a clear-cut operational distinction between the general context and other more specific cues to the response

Regardless of the theoretical interpretation, the parametric studies of Peterson and his associates indicate that retention of single items is a function of the same task variables that are known to influence the acquisition and retention of lists, e g , meaningfulness and distribution of practice The same appears to be true for the retention of a single item within a list Using lists of word number pairs, Peterson, Saltzman, Hillner, and Land (1962) found that the probability of recall of individual associations decreased steadily as a function of the time interval (or number of items intervening) between presentation and test Both the presentation and test occurred during a continuous exposure of the list, and some of the retention intervals were shorter than the intertrial intervals conventionally used in the learning of lists Forgetting was significantly less after two than after one presentation of an item Since parallel functional relations appear to apply to individual items and to lists, there is every reason to agree with the conclusion of Peterson et al that "it would seem parsimonious to assume that short term retention is basically similar to retention over longer intervals until experimental evidence dictates otherwise" (1962, p 402)

Short-Term Retention and the Concept of Memory Trace

The evidence reviewed thus far in this section clearly supports the conclusion that substantial amounts of forgetting may occur over short periods of time. A variety of experimental procedures has been used to demonstrate such rapid forgetting, most of which have certain basic features in common. (*a*) There is only one exposure to the materials. (*b*) The retention interval is filled with an interpolated activity which is assumed to prevent rehearsal. This interpolated activity consists either of the recall of other items or of exposure to irrelevant stimuli. In a broad sense, therefore, the experimental procedures usually reduce to studies of short term retroactive inhibition. (*c*) With few exceptions, ordered recall of the entire stimulus series is used to measure retention.

The fact of rapid short term forgetting in the absence of rehearsal has been widely interpreted as lending at least indirect support to the concept of an immediate memory trace which deteriorates rapidly unless it is restored and converted into a more permanent trace by repetition. The question may again be raised of whether the concept of a decaying trace and the implied distinction between short term and long term memory (Hebb, 1949, Broadbent, 1958) are indeed supported by the data.

The finding from which the concept of a decaying memory trace appears to derive most of its face validity is the extreme speed with which performance deteriorates under the experimental conditions described above. The retention curve may in a matter of seconds drop to a level comparable to that obtained after days or even weeks in conventional studies of rote learning. Thus, a dual process is suggested, one reflecting the fate of the primary trace and the other that of a consolidated trace which has been prevented from decaying by explicit or implicit rehearsal. It is uncertain, however, whether the results obtained in experiments on immediate memory and on long term memory are sufficiently comparable to make a judgment about the relative rates of forgetting in the two situations and then to infer the operation of two different processes of retention.

Two variables which are known to determine the rate of forgetting are the degree of original learning and the measure of retention used. It is not possible to compare directly the degrees of learning represented by a single exposure to a short series of discrete units on the one hand and the extended practice of a list on the other. The measure of retention used in many of the studies of short term retention, e g , those which use materials such as series of digits or letters, is perfect reproduction of the entire sequence. In studies of long term retention credit is typically given for each item reproduced correctly. When these differences are taken into account, the possibility remains open that there is no basic discontinuity between the

retention functions obtained in the two types of experimental situation The evidence showing the retention of single items to be a function of the same variables as the acquisition and retention of lists strongly supports the argument for continuity

A more fundamental question concerns the possibility of drawing inferences about the properties of a hypothetical memory trace on the basis of the experimental observations which have been made As we have emphasized, one of the characteristics of the relevant studies is the use of interpolated activities (or such special arrangements as simultaneous stimulation of the two ears) for purposes of preventing rehearsal The amount of forgetting is, therefore, determined jointly by the retention interval and the effects of the interpolated activity The degree of decay of the trace must, of course, be assumed to be a function of the time interval per se The assumption has to be made that the interpolated activities are not primarily responsible for the retention losses and do not influence the rate of decay of the trace It should be noted that interpolated materials of varying degrees of similarity to the original learning materials have been used When the interpolated activity involves the recall of other items from the same list, as in experiments in which length of retention interval is co ordinated with order of recall, the similarity is high In other situations, e g , when the interval after presentation of a verbal item is filled with counting backwards, the similarity is low

It is clearly hazardous to assume that interpolated activities merely serve to prevent rehearsal and do not function as effective sources of interference over short retention intervals The objection to this assumption is not removed when the learning materials and the interpolated stimuli are highly dissimilar In conventional rote learning studies substantial amounts of retroactive inhibition have been obtained even when the intertask similarity was low Such interference appears to be largely a matter of "generalized response competition," i e , the *S*'s tendency to persist in the performance of the interpolated task when required to recall the original list (Newton & Wickens, 1956, Postman & Riley, 1959) Performance decrements owing to a loss of set are very likely to occur in experiments on short term retention in which *S*s are required to switch rapidly from one activity to another The highly disruptive effects which can be produced by brief interpolated tasks filling only a fraction of the retention interval are illustrated by Conrad s (1960a) findings discussed above The fact that the kinds of activities which are used to fill the retention intervals also produce significant proactive effects likewise points to the presence of interserial interferences It has already been noted that short term proactive interference cannot be readily interpreted from the point of view of decay theory In summary, experimental tests of the decay hypothesis are

seriously, if not fatally, handicapped by the dual requirement of measuring forgetting as a function of time and preventing rehearsal The latter requirement makes it necessary to fill the retention interval as effectively as possible with an interpolated activity Once that is done, the relationship between forgetting and the sheer passage of time is obscured

At a still more general level, the connotations of the concept of *trace* continue to suffer from vagueness The assertion that each stimulus leaves some neural aftereffect has few, if any, precise empirical implications If the trace is to be specified in more detail, the extremely troublesome question arises of what the unit of stimulation is which is represented by a fading trace is it an individual element such as a letter, a group of elements such as a word, the entire series? When the functional unit changes, as it appears to do when 'chunking' takes place in the immediate memory span, or in the recall of individual items, does the trace unit change likewise? As long as such questions remain unanswered, it is not likely that critical tests of the trace hypothesis can be devised

The recent upsurge of interest in the decay theory of memory has led to the development of some very valuable new procedures for the measurement of short term retention Important new information about the course of retention over short intervals has been obtained by these procedures However, the theoretical clarification of the concept of trace continues to lag far behind these experimental developments

Free Recall and Serial Recall

Associative processes and preexperimental habits of classifying and ordering the learning materials are likely to be aroused whenever a series of items is recalled There is only limited opportunity to evaluate the influence of such habits when the series consists of homogeneous units, and reproduction in the serial order of presentation is required, as in the determination of the memory span and in many of the studies of short term retention which have been reviewed thus far It is true that preexperimental habits can influence the grouping or "chunking' of items and are reflected in the characteristics of the *S*'s errors In general, however, homogeneity of the units in the list and the constraints of serial order clearly limit the possibilities of 'recoding' and rearranging the learning materials The opportunity to observe and evaluate the transfer effects of preexperimental habits and the operation of associative mechanisms in the reproduction of verbal items are greatly enhanced when the method of free recall is used with materials which exceed the span of immediate memory We turn now to studies of short term retention by the method of free recall and to an evaluation of the differences between free recall and serial recall

Experimental procedures.—When a subject has been exposed to a list of items, he may be instructed to reproduce them either in the order of presentation or in any order he wishes. These instructions define serial recall and free recall, respectively. When the series is presented more than once, there are three combinations of conditions defined by the constraints on serial order in presentation and recall: (*a*) constant serial order of presentation and serial recall, (*b*) constant serial order of presentation and free recall, (*c*) varying order of presentation and free recall (Waugh, 1961). The alternation of presentations and recalls defines a rote learning task in which learning is measured by the rate of change in the amounts retained as a function of the number of repetitions (Murdock, 1960). An important difference between these procedures and other rote learning tasks is that study trials are separated from test trials. Such a separation can also be made in paired associate learning—as, for example, in many recent studies of one trial learning (e.g., Clark, Lansford, & Dallenbach, 1960; Estes, Hopkins, & Crothers, 1960; Postman, 1962a; Rock, 1957; Underwood, Rehula, & Keppel, 1962).[3] The separation of study trials and test trials is, of course, impossible in serial anticipation learning.

As methods of measuring performance, serial recall and free recall occupy positions intermediate between the memory span experiment on the one hand, and conventional rote learning procedures on the other. The methods of free recall and serial recall, singly and in juxtaposition, permit the analysis of a number of important relationships which cannot be investigated directly in other task situations: (*a*) the effects on retention of serial position in the list as distinct from the effects of serial constraints in recall, (*b*) the relationship between the order of recall and the frequency of recall, (*c*) the relationship between serial position in the list and order of recall, (*d*) the effects of interitem associations on the frequency and order of recall. Questions such as these become highly important in investigations of the effects of language habits and associative hierarchies on learning and retention. It is primarily in such studies that the recall methods, and especially that of free recall, have received their most effective systematic application.

Basic task variables in free recall.—Recent analyses of the variables determining performance in immediate free recall have provided important information about the processes which come into play when the span of immediate memory is exceeded and the constraints of serial recall are re-

[3] Under the conventional method of anticipation the *S* receives immediate reinforcement or correction. When study trials and test trials are alternated, there is only delayed feedback. Nevertheless, the rate of acquisition under the alternation procedure equals or surpasses that by the conventional method (Battig, 1961; Lockhead, 1962).

moved The most general conclusion which is supported by the evidence is that amount of recall depends on (*a*) the opportunities for rehearsal, (*b*) the availability of integrated response units, and (*c*) the number and strength of the interitem associations which exist or can be established within the series The effects of three major task variables in free recall—length of list, time of presentation per item, and meaningfulness—can be interpreted as consistent with this generalization.

In an extensive investigation of the variables determining the immediate free recall of unrelated words Murdock (1960) has presented clear evidence for the conclusion that the number of items recalled after one presentation of a series is a linear function of the total time required for presentation The total time is the product of the length of the series and the presentation time per item The amount recalled remains invariant as long as this product does Recall increases as a function of both length and presentation time, but a reduction in one of these can be compensated for by an increase in the other In Murdock's data this invariance holds true over a wide range of the two variables and of experimental conditions, and he showed it to be present in the results of other investigators as well It appears likely, as Deese (1960) has suggested, that increases in either the length of the list or the presentation time per item enhance the opportunity for rehearsal To account for the invariance of amount recalled with time, it is necessary to assume further that only a limited number of items can be rehearsed during a given period of time, regardless of how many different items are presented Rehearsal will strengthen the connection between each item and the situational context, but it will serve other functions as well, depending on the characteristics of the items in the list Among these characteristics meaningfulness is of major importance

In the analysis of rote learning it has proved useful to distinguish between two stages in the process of acquisition a response-learning stage and an associative stage (Underwood & Schulz, 1960) During the response-learning stage the items to be recalled become available as responses If they are not already in the *S*'s repertoire, they must be integrated through practice, if they are already available as integrated units, the range of responses must be restricted to those in the list. During the associative stage the prescribed sequential connections between cues and responses are established It is during the response-learning stage that meaningfulness has its major effects by determining the order in which responses become available for association When free recall is considered within the framework of this analysis, it becomes apparent that response integration is just as essential here as in other situations in which recall of the items is required However, response learning, in the narrower sense of restriction of the repertoire to the required units, now becomes closely tied to as-

sociative learning Response learning and associative learning are interdependent in the free-recall situation because there are no restrictions on the order of recall Thus, any and all associations which exist or are established among the integrated and available items will facilitate reproduction By contrast, in paired associate and serial learning, all but the one required association become sources of intraserial interference It follows from these considerations that amount of free recall should be a function of (*a*) degree of integration of the response units and (*b*) strength of interitem associations when the list is composed of integrated units The facts are in accord with these conclusions

Under otherwise comparable conditions, free recall for meaningful words is very much higher than for nonsense syllables (Postman & Adams, 1956b, 1958, Postman & Phillips, 1961) This difference increases with the length of the list For nonsense syllables, the amount of free recall is a function of association value (Postman, Adams, & Phillips, 1955) Since association value and degree of response integration are closely related (Underwood & Schulz, 1960), all these results support the critical importance of response integration in free recall The interaction between meaningfulness and length of list in these comparisons indicates that rehearsal serves different functions for the two type of materials enhancing response integration for nonsense material and associative learning for meaningful items It is reasonable to suppose that the former process develops more slowly as a function of rehearsal time than the latter

When the learning materials are restricted to integrated units, the probability of interitem associations becomes decisive This conclusion is clearly supported by the results of a study by Deese (1960) investigating the effects of the frequency of usage of words and length of list on the amount of free recall A wide range of word frequencies was compared, with the words selected at random within each frequency class, and lists varying in length between 12 and 100 items In agreement with earlier studies (Bousfield & Cohen, 1955, Hall, 1954) recall was found to increase with word frequency, and there was also a positive relationship between length of list and amount recalled The increase as a function of length was significantly greater for lists of high than of low word frequency Deese s analysis shows that both the main effect of word frequency and its interaction with length of list can be attributed to correlated variations in the strength of interword associations An independent index of associative structure (the average frequency with which the words in a given list elicited every other word in the list as an associate) correlated with both word frequency and recall With the index held constant at zero, there was no residual effect of word frequency on recall The index also increased with length of list but more so for lists of high than of low word frequency

This interaction accounts for the disproportionately large increases in recall as a function of length found for the high frequency lists Deese suggests that these increases represent a joint effect of associative clustering and rehearsal

There is other clear evidence for the dependence of free recall on the strength of interitem associations Instead of determining the amount of interword association in randomly selected lists, it is possible to manipulate this variable systematically by constructing series for which the index of associative structure varies from low to high When that is done (Deese, 1959a), a high correlation between the strength of interitem associations and amount recalled is obtained At the same time, a negative correlation is obtained between the index and the frequency of intrusions from outside the list the higher the index, the more likely it is that associations aroused by a given item are actually members of the list. Finally, the higher the index the more agreement there is among *Ss* on the intrusions that do occur, i e, the context of the list determines the particular intrusions which are given When the items in a series converge upon a common associative response outside the list, the frequency with which that response will occur as an intrusion can be predicted reliably on the basis of the converging associative strength (Deese, 1959b)

Given integrated response units, the average strength of the associations linking the items in a list is a powerful determinant of the amount recalled The question remains of how interitem associations function during the actual process of recall There is reason to believe that *Ss* do not deliberately reconstruct the lists on the basis of the associative context For example, when *Ss* are provided with an extra list cue labeling the associative context of the list, there is no increase in the amount recalled (Deese, 1959a) Recently Rothkopf and Coke (1961) have shown that after a single exposure to a very long list (all but one of the stimulus words in the Kent Rosanoff list) the probability of recall of a given word varies directly with the number of other items in the list which elicit that word as an associative response There is a large number of different associative groupings within such a list, yet recall remains highly sensitive to the convergence of associative probabilities Associates appear to elicit each other directly, regardless of contiguity during exposure and not through a process of deliberate reconstruction It follows that the order in which responses are emitted should be governed by the preexperimental connections among them The facts of clustering in recall show that such is indeed the case

Clustering in recall—It will be useful to make explicit and to emphasize the distinction which has just been made between the effects of preexperimental habits and associations on the amount retained on the one

hand, and on the performance characteristics of recall on the other. Consideration of the performance characteristics provides an independent check of hypotheses advanced to account for the observed effects on retention. Associative clustering is a case in point. If interitem association enhances recall because items elicit each other just as they do in free association (Deese, 1961), it is clear that associated pairs should appear together in recall. The degree of clustering, moreover, should be a function of associative strength. The facts support these predictions. When members of associated pairs are presented in scrambled order, they tend to be grouped together in recall (Jenkins & Russell, 1952), and the amount of clustering is a function of both the rank of the response word in the free association norms (Rothkopf & Coke, 1961, Russell & Jenkins, 1952) and the absolute frequency with which the response appears in the norms (Jenkins, Mink, & Russell, 1958). Since degree of clustering and amount of recall are correlated, the data are consistent with the assumption that gains in recall occur when associates elicit each other directly. It should be noted, however, that the amount of clustering can be substantially increased if *Ss* are instructed at the time of the recall test to reproduce as many pairs as they can (Postman, Adams, & Bohm, 1956). Interestingly enough, the increase in clustering under instructions is not accompanied by increases in the total amount recalled. This finding supports the view that deliberate selection is not responsible for the gains in recall as a function of associative probability.

The story does not, however, end here. There is another important kind of clustering which has been traditionally attributed to the operation of mediating processes. We refer here to category clustering which has been studied extensively by Bousfield and his associates. The categories used in these studies are taxonomic, e.g., animals, vegetables, professions, names, etc. The basic experimental procedure (Bousfield, 1953) involves the presentation in random order of a series of words falling into a limited number of categories, followed by a test of free recall. The frequency with which sequences of related words occur in recall defines the degree of clustering. The degree of clustering has been shown to increase as a function of several variables: (*a*) the number of presentations of the list (Bousfield & Cohen, 1953), (*b*) the frequency of usage of the words in the list (Bousfield & Cohen, 1955), (*c*) the frequency of occurrence of the words in the cultural norms for categories (which is correlated with word frequency) (Bousfield, Cohen, & Whitmarsh, 1958), (*d*) the number of categories in the list (Bousfield & Cohen, 1956). In general, amount recalled is correlated positively with the degree of clustering although the relationship is somewhat unstable. For example, the findings on amount recalled as a function of the number of categories are inconsistent (Bous-

field & Cohen, 1956, Mathews, 1954) During the recall period clustering characteristically rises to a maximum and then declines as a function of time As the degree of clustering increases, the initial level becomes higher and the maximum rate is reached sooner

In his theoretical interpretation Bousfield has suggested that exposure to individual words results in the activation of superordinate (neural) structures which represent the category to which the word belongs Arousal of a superordinate structure in turn facilitates the response of subordinate structures related to it, i e , the emission of other words in the category These organizational systems, which are modeled after Hebb's cell assemblies, are assumed to have the properties of habits Considered as a habit, a word is, therefore, assumed to have two kinds of strength—habit strength reflecting its frequency of usage, and a "relatedness increment," which is attributed to the facilitation produced by activation of the superordinate It is the relatedness increment which is responsible for clustering and correlated increases in recall The relatedness increments are assumed to develop during both learning and recall

To the extent that the words elicit each other as associates, the assumption of a mediating process is not required, and category clustering can be considered a special case of associative clustering Clustering may, however, be based on formal grammatical as well as semantic categories (Cofer, 1959, Gonzalez & Cofer, 1959) In such cases, the simplest assumption is that a common cue producing response given to each item in a category mediates sequential reproduction Such an interpretation of the mechanism of clustering was, indeed, recently suggested by Bousfield, Whitmarsh, and Berkowitz (1960) who showed that there is a significant correlation between incidence of clustering and what they describe as 'partial response identities," i e , the extent to which a set of words elicits *common* associative responses A response mediated mechanism is also indicated by the finding that both instructions about the principle of construction of the list as well as practice in the recall of a sample list produced significant increases in the amount of clustering (Bousfield, 1955)

The relationship between the two types of clustering remains uncertain Associative clustering appears to reflect directly the associative structure of the words in the list, and no recourse to a mediational process is compelled by the data On the other hand, at least some forms of category clustering strongly suggest mediation by cue producing responses The question now is whether it is necessary to assume two distinct types of clustering and if it is, what their relative priorities are in determining sequential order in recall Does mediated clustering supersede or supplement associative clustering once the presence of taxonomically related groups of items is recognized and appropriate cue producing responses are

made to the words? To answer this question, it will be necessary to develop materials in which the effects of interitem association and response-produced mediators are pitted against each other

Sequential dependencies in free recall—The question of the specific mechanism by which nonrandom relations among words influence recall becomes critical in the interpretation of the effects of statistical dependencies between verbal items The method of statistical approximations to English introduced by Miller and Selfridge (1950) has provided a method for the systematic manipulation of the degree of contextual determination of the successive items in a list In continuous discourse the range of possible alternative responses decreases with the length of the sequence providing the context of successive items By varying the length of the contextual sequence determining the choice of successive words, Miller and Selfridge constructed lists which represented different orders of approximation to the statistical structure of English Number of word retained was a negatively accelerated function of the order of approximation (cf also Postman & Adams, 1960, Sharp, 1958) Thus, the gains in recall appeared to be attributable to short term dependencies rather than to the meaning of the entire passage [4] However, when the immediate memory span method is used so that recall in the exact order of presentation is required, the function is positively accelerated and textual material has a distinct advantage (Marks & Jack, 1952) The more sensitive the test of retention is to the preservation of sequential dependencies, the greater are the effects of contextual constraints

Why does recall increase with the order of approximation? The gains cannot be determined by the strength of interitem associations since sequential dependencies in continuous discourse clearly cannot be derived from the associative hierarchies of individual words, especially when the critical importance of function words is considered An analysis by Deese (1961) suggests that increases in the order of approximation influence not so much the amount retained as the *S*'s ability to reconstruct the material, essentially by guessing, on the basis of his knowledge of the characteristics of the language One persuasive bit of evidence presented by Deese is that the gains in recall as a function of order of approximation are highly correlated with the degree of agreement among *S*s in supplying words deleted from the different passages Words on which there is high agreement are

[4] A recent analysis by Coleman (1963) shows that the higher-order approximations included more complex words and were grammatically more awkward than the lower order ones When passages were matched in syllabic length and word frequency, prose was recalled better than higher-order approximations. Coleman also showed that the differences among orders of approximation depend on the length of the sequence used as the unit of recall The longer the sequence the greater is the advantage of the higher order approximation

likely to be the ones which appear in the original passage It is reasonable to suppose that *Ss* do in recall what they do on the deletion test, viz, reconstruct the material on the basis of the probabilities of different items and sequences Another experimental finding is consistent with the view that recall of the higher-order lists is in part constructive in nature When *Ss* are required to recall their own responses in a deletion test, recall first increases and then decreases as a function of the commonality of the responses This relationship obtains under both intentional and incidental conditions (Postman & Adams, 1960) Thus, *individual* words on which there is a high degree of contextual constraint are recalled no better than words on which there is little agreement High-commonality responses are likely to be parts of conventional phrases or function words which are determined mostly by the context of the immediately preceding words rather than by the general context Such locally determined responses, which occur in many different verbal environments, are not likely to be recovered in free recall in the absence of the local context. These very items can, however, be readily reconstructed in the reproduction of a sequence and contribute to the gain in recall as a function of the order of approximation.

Recognition of the constructive nature of the recall of sequentially ordered series points up an important methodological problem which arises in the evaluation of free recall protocols for meaningful words and connected discourse The problem concerns the appropriate baseline to be used in measuring the amount recalled If this measure is to reflect changes in performance attributable directly to the presentation of the list, the level of accuracy which can be achieved by guessing must be taken into account The question becomes how much that is specific to the list must be recalled in order to allow reconstruction of the rest The problem arises whenever the series to be learned includes redundant sequences, and the learner has been taught to make use of these redundancies (Aborn & Rubenstein, 1952, Rubenstein & Aborn, 1954) or has an opportunity to discover them during practice on the learning materials (Miller, 1958)

Serial order in recall —Analysis of the effects of sequential dependencies has led to important advances in the understanding of the conditions determining the serial position curve in free recall. In contrast to the typical curve obtained in anticipation learning, the serial position curve in free recall shows a substantially greater finality than primacy effect Since the advantage of the terminal positions is enhanced when rehearsal is prevented, it had been suggested that the primacy effect is produced by rehearsal of the initial items during presentation of the list (Welch & Burnett, 1924 Raffel 1936) While rehearsal probably has some effects (cf Bousfield Whitmarsh & Esterson, 1958), it is now clear that the shape of the

serial position curve in free recall is determined in large measure by the order of emission of the items. The order of emission can, in turn, be manipulated by varying the degree of sequential dependency between successive items.

These relationships were demonstrated in a study by Deese and Kaufman (1957). When the words in the list are unrelated, there is a high degree of correlation between the probability of recall and order of recall. This may be regarded as an extension of Marbe's law to free recall (Bousfield, Cohen, & Silva, 1956). The last items are recalled most frequently (perhaps because the retention interval for them is shortest) and are given first. Thus, the finality effect is greater than the primacy effect. When successively higher orders of approximation to English are used, the order of recall comes to be determined by the order of the items in the list, and relative primacy increases. Priority in recall is favorable to retention, and the beneficial effects of early emission counteract the influence of recency as such. As a result, the serial position curves in the free recall of textual material and in serial anticipation learning are quite similar. Analogous shifts in serial position can also be produced by instructions for serial recall (Deese, 1957). In all these cases sequential constraints, either inherent in the material or imposed by instructions, determine the order of recall and thereby the shape of the serial position curve.

The fact that order of recall per se, as distinct from order of presentation, influences performance in an unpaced test was demonstrated in a study by Waugh (1961) in which three conditions of acquisition were compared: (*a*) serial recall, (*b*) free recall with varying orders of presentation, and (*c*) free recall with a constant order of presentation. A linear learning curve was obtained by the method of serial recall, with successive increments on the order of the immediate memory span. On the other hand, negatively accelerated learning curves, which were virtually identical, were obtained under the other two conditions. Thus, the instructions concerning the order of recall rather than order of presentation was the critical variable in determining the course of acquisition. Waugh suggests that under conditions of serial recall *S*s rehearse the list in segments falling within the memory span while effectively ignoring the rest of the list. This is a strategy reminiscent of that described in connection with the running memory span. Such a strategy is not adopted when there is no restriction on the order of recall, and the probability of new words being recalled decreases with trials. The slow rate of change in the amount recalled would be consistent with the presence of a steep finality effect.

Conclusion: free recall versus ordered recall.—The evidence reviewed in this section justifies the conclusion that the distinction between free recall and ordered recall is of major significance in the classification of studies

of retention in general, and short term retention in particular The presentation of verbal materials is necessarily sequential, and the changes produced by practice are of two kinds—in the availability of the correct responses and in the ability to order these responses The separate assessment of these two components of acquisition becomes possible by comparisons between free recall and ordered recall

The functional relations obtained by the method of free recall will diverge from those obtained by the methods of ordered recall to the extent that the experimeotal variables have differential effects on response availability and on the retention of order The influence of response similarity is a case in point. Amouot of free recall varies directly with response similarity as may be expected on the basis of the positive correlation between degree of category clustering and recall By contrast, interitem similarity is inversely related to the accuracy of ordered reproduction These opposing trends have been clearly demonstrated in experiments making direct comparisons between the effects of response similarity on free recall on the one hand, and on reconstruction of serial order (Horowitz, 1961) and paired associate learning (Underwood, Runquist, & Schulz, 1959) on the other It is clear that the variables determining short term retention must be considered in relation to the method of measurement Any general theory of retention, whether a trace hypothesis or an interference hypothesis, must take account of this fact

Incidental Learning

The consideration of incidental learning has been reserved for the final section of this paper because much of what has been reviewed above is propaedeutic to this discussion The operations used in studies of incidental learning are typically those of experiments on the short term retention for materials exceeding the memory span The strategy of research has been to examine under incidental conditions functional relations known to obtain in intentional learning The question which has typically been asked is whether and in what ways recall changes when instructions to learn are omitted This approach may be characterized as a "subtraction method," and the evaluation of the experimental findings must take its departure from the known facts and theoretical interpretations of intentional learning

For a long time experimental investigations were limited to the sampling of a variety of materials in order to demonstrate that incidental learning does, indeed occur and is significantly inferior to intentional learning (e g, Myers, 1913, Shellow 1923) Considerable attention was given to changes as a function of age and other psychometric variables The theoretical analysis of incidental learning and the systematic analysis of its determi

nants have been developing slowly This neglect of the problem was due in large measure to the difficulties which arise in the formulation of a clear operational distinction between intentional and incidental learning

Definitions and Experimental Designs

Definition of incidental learning.—A basic obstacle to the definition of the term "incidental" has been that its connotations are negative, i e , it refers to the *absence* of a set or intent to learn This negative definition can, to be sure, be given operational meaning by omitting instructions to learn But in then assuming that learning under these conditions occurs without any intent to learn, one is essentially in the position of accepting the null hypothesis, and there has been considerable reluctance to do so Thus, McGeoch wrote " much of the learning which goes on with no overt instructions is, nonetheless, influenced by implicit instructions and sets certainly it cannot be said with any conclusiveness that there are experiments in which implicit sets have not operated, but, more than this, probability is on the side of the hypothesis that all of the results [in incidental learning] have been determined by set' (1942, p 304)

McGeoch's point was that it is hazardous ever to assert that learning is incidental in an absolute sense We can accept this point without abandoning the substantive problem which is implied in the distinction between intentional and incidental learning Instead of seeking to demonstrate a dichotomy, we must shift our concern to the functional relations between the instruction stimulus on the one hand and measures of learning and retention on the other The instruction stimulus is an integral part of the conditions which must be specified in any investigation of learning It can also be manipulated systematically, and one of the dimensions along which it can vary is the amount of information given to the *S* about the test of performance which he is to expect

When the instructions do not prepare the *S* for a test on a given type of materials, it is convenient to designate the learning of these materials as incidental This designation should not imply that such learning occurred in the absence of any incipient or transitory sets Whether or not such sets are likely to have been aroused becomes a matter of theoretical interpretation One may, in fact, wish to conceptualize the effects of instructions on learning in terms of the strengths and specificity of the sets which are aroused In the first instance it is necessary, however, to separate the problem of operational definition from theoretical interpretation Operationally, incidental and intentional learning are distinguished by the use of different classes of instruction stimuli—those which do and those which do not prepare the *S* for a test of retention In practice, manipulation of the instruction stimulus is often supplemented by a postexperimental inquiry which

ascertains the *S*'s response to the instructions Incidental *S*s who anticipated a test or deliberately rehearsed the material are discarded and replaced Again, the verbal reports should not be taken as absolutely valid, but the screening of *S*s on the basis of their statements sharpens the separation between the experimental conditions Similarly, in an experiment on pitch discrimination one would eliminate *S*s who admitted to having made judgments of loudness rather than pitch In summary, then, the terms "incidental" and "intentional" are defined by categories of instruction stimuli, supplemented by a classification of *S*s on the basis of their verbal statements It is not necessarily assumed that these definitions imply a sharp discontinuity with respect to the operation of sets or other intervening processes

Experimental designs—Two types of incidental learning situations may be distinguished (cf Kausler & Trapp, 1960, Mechanic, 1962a, Postman, Adams, & Bohm, 1956, Postman & Senders, 1946) In Type I the *S* is exposed to the stimulus materials but given no instructions to learn Following the exposure his retention is tested unexpectedly The choice of the test determines the criteria of incidental learning in a given experiment These criteria may vary with respect to both the kind and amount of learning that is required for successful performance, depending on whether retention is tested by recognition, free recall, serial recall, or transfer to a new task In the interpretation of experimental findings it is important to bear in mind that conclusions about incidental learning are specific to the method of measurement, just as they are in the evaluation of intentional learning While this point is obvious, it has not always been heeded in comparisons among experiments on incidental learning

In Type II of incidental learning the *S* is given a specific learning task but during practice is also exposed to materials or cues which are not covered by the learning instructions His retention for those features of the situation which are not relevant to the task specified in the original instructions defines the amount of incidental learning, and the measure obtained will again be a function of the test Type II situations may be further subdivided into two classes on the basis of the relationship between the relevant and irrelevant components of the total learning situation The irrelevant components may be features or attributes of the materials which the *S* has been instructed to learn but which are irrelevant in the sense that their discrimination and retention are not required for the performance of the task defined by the experimenter For example, if the verbal items which *S* has been instructed to learn are printed in different colors, the colors are a feature of the learning materials which is irrelevant to the explicit task On the other hand, the irrelevant components may be materials or cues which bear no direct relation to the learning task, e g , when

the instructions are to learn a series of words but such additional items as digits or geometric forms are exposed along with the words Thus, the two classes which are distinguished within the Type II situation refer respectively to the incidental learning of intrinsic and extrinsic components of the experimenter-defined learning task

Type I and Type II are both incidental-learning situations because the *S* is not instructed to learn the materials on which he is tested In both situations the basic question is whether these materials elicit responses which become associated with the experimental situation and which can then mediate correct performance on the test of retention At the same time, however, there are important differences between the two kinds of experimental arrangements These differences derive from the fact that instructions to learn are given in Type II but not in Type I The learning instructions, however circumscribed and specific they may be, must be assumed to have two consequences (a) The general class of responses which is entailed by a set to learn, such as rehearsal and the differential labeling and categorizing of items, is activated To the extent that such instrumental responses generalize to the irrelevant features and materials, the Type II situation is in principle more favorable to incidental learning than is the Type I situation In the former, cue producing responses are generalized to the irrelevant items whereas in the latter they must be aroused by these materials alone (*b*) In the Type II situation the critical differential responses to the irrelevant materials must be given along with those to the relevant materials When the exposure interval is limited, the responses to the relevant and irrelevant materials are, therefore, in competition with each other (Mechanic, 1962a) No such task competition obtains in the Type I situation In that sense the Type II situation is less favorable to incidental learning than the Type I situation The net difference between the two situations will depend on the extent to which the beneficial effects of a generalized set to learn are offset by the effects of task competition

Which of the two experimental arrangements is more appropriate depends on the theoretical question at issue When the emphasis is on associative processes in incidental learning as determined by the nature of the materials and the conditions of presentation, Type I is favored because an evaluation of these variables can be made directly without the complications introduced by generalization of sets and task competition On the other hand, investigators whose primary concern is with incidental learning as a function of motive and incentive conditions have usually turned to the Type II situation Such investigations have focused on the question of how generalization of set and task competition are influenced by variations in drive and incentive However, the difference between Type I and

Type II must not be overstated While it is convenient to make this distinction for purposes of classifying experiments, it is also important to recognize the continuity between them This point will become apparent when we consider the role of the orienting task in incidental learning

Orienting Tasks

In a study of incidental learning the experimental arrangement must be such as to ensure the exposure of the *S* to the learning materials The particular procedure used for this purpose constitutes the orienting task To be useful, an orienting task must satisfy two criteria (*a*) it must, indeed, create conditions which make it certain that the *S* perceives the incidental stimulus materials, and (*b*) it should minimize the development of uncontrolled sets to learn The requirement of an orienting task raises important problems of control and interpretation in the experimental analysis of incidental learning

Type I —Consider the Type I situation first In some of the early studies using this design (e g , Biel & Force, 1943, Mulhall, 1915, Porter, 1938), the incidental *Ss* performed an orienting task whereas the intentional *Ss* did not but were simply instructed to learn Under these conditions it is impossible to determine the degree to which intent to learn and freedom from interference by the orienting task contribute to the advantage of the intentional learners To separate these two effects, a three-group design is required two intentional groups, one of which does and one of which does not perform the orienting task, and the incidental group which likewise performs the orienting task When this design is used, it is typically found that the performance of the orienting task reduces the amount learned under intentional conditions, as a consequence, the difference attributable to intent per se is correspondingly reduced (Gleitman & Gillett, 1957, Postman & Adams, 1956b, Postman, Adams, & Bohm, 1956, Saltzman, 1953)

The activities required by an orienting task may be more or less consistent with the responses which mediate associative learning One may conceive of a continuum of orienting tasks, ranging from those requiring responses maximally favorable to learning to those requiring responses maximally antagonistic to learning For example, when the task is to give meaningful associations to each of a series of items, the responses are similar to those which are assumed to occur in learning By contrast, guessing one of a set of numbers to each of a series of verbal items or matching such items to a limited set of geometric designs is probably unfavorable to learning because the repeated use of the same response is likely to result in response-produced generalization

When the orienting task is performed by both groups, the difference

between intentional and incidental learners should depend on the position of the orienting task along the continuum described above. It should be minimal when the orienting task falls at either extreme of the continuum. At the unfavorable extreme, the beneficial effects of intent would be minimized by massive interferences from the orienting activity while at the favorable extreme there would be maximal facilitation of incidental learning. The data support these expectations. With materials and conditions of testing held constant, the differences between intentional and incidental learners may vary over a wide range as a function of the orienting task (Postman & Adams, 1956b, Saltzman, 1956). What is most important, however, is that the difference reduces to zero under conditions which appear to approximate the two extremes of the continuum of orienting tasks, i.e., when learning is either seriously hindered or substantially facilitated by the orienting task. These facts lead to the conclusion that intent per se has no significant effects on learning. All its effects are indirect, i.e., instructions to learn activate *responses to the materials* which are favorable to acquisition. The same results can be achieved by appropriate orienting tasks without instructions to learn.

Type II.—In the Type II situation, the instructions to learn the relevant materials at the same time impose an orienting task for the irrelevant materials, provided both components of the task are intrinsically associated with each other. These conditions are exemplified by the study of Bahrick (1954) in which *S*s were instructed to memorize a series of geometric forms and were later tested for retention of the colors filling each of the forms (cf. also Kausler, Trapp, & Brewer, 1959). However, when the irrelevant materials are extrinsic to the relevant ones, an additional orienting task is required just as in the Type I situation. An experimental procedure used by Mechanic (1962a) illustrates this situation. Pairs of trigrams were presented, one member above the other. The *S*s were instructed to learn one member of each pair, e.g., the bottom one. To ensure exposure to the irrelevant item, *S*s were also instructed to rate the members of each pair for phonetic similarity. As Mechanic has pointed out, the procedure in which relevant and irrelevant items are intrinsically associated permits no control over the *S*'s response to the incidental material even though its sensory reception is ensured. When the irrelevant items are extrinsic to the relevant ones, the response to the incidental materials can be manipulated by means of different orienting tasks.

The systematic difference between these two classes of Type II situations is brought out by the divergent effects of strength of incentive under these procedures. In the study mentioned above, Bahrick (1954) found that the amount of incidental learning was inversely related to the strength of the incentive for performance of the intentional task. Amount of intentional

learning, on the other hand, increased directly with the strength of the incentive Thus, both intentional and incidental learning varied as a function of incentive, but in opposite directions Similar results had previously been reported by Bahrick, Fitts, and Rankin (1952) concerning the effects of incentive on the performance of central and peripheral tasks By contrast, under the conditions of Mechanic's experiment variations in incentive influenced neither intentional nor incidental learning (Mechanic, 1962b) These contradictory findings fall into place when the differences in the degree of control over the *S*'s responses in the two situations are taken into account It is reasonable to suppose that in a situation like Bahrick's variations in incentive influenced the relative frequency of differential responses to the intentional and incidental materials and as a consequence produced significant shifts in the amounts of intentional and incidental learning When responses to the materials are controlled by an orienting task, incentives do not influence either type of learning On the other hand, as Mechanic was also able to show, variations in the orienting task determine the amount of Type II incidental learning in a manner consistent with that found in Type I situations

Variations in drive, like incentive, may be expected to influence the relative frequency of the *S*'s responses to relevant and irrelevant stimuli Several studies have investigated Type II incidental learning as a function of anxiety manipulated either by threat of punishment or by the selection of *S*s on the basis of test scores The results are not consistent, some experiments showing systematic effects on incidental learning (Silverman, 1954, Silverman & Blitz, 1956), others failing to obtain an effect (Kausler, Trapp, & Brewer, 1959) This lack of consistency is perhaps not surprising since it cannot be predicted with certainty exactly how the distribution of responses to the relevant and irrelevant stimuli will be influenced by anxiety in a given situation The more directly these responses are brought under experimental control, as by the manipulation of the orienting task, the more likely it is that the relative amounts of intentional and incidental learning can be altered in predictable ways

Functional Relations in Incidental Learning

Analysis of the effects of the orienting task has led us to the conclusion that intent per se is not a significant variable in learning but that the instruction stimulus influences the amount and characteristics of learning by determining the differential cue producing responses, including deliberate rehearsal, which occur during the period of practice These differential responses include categorizing responses such as naming or labeling, other responses elicited by the items through stimulus generalization, and responses serving as associative links among the members of a series It is the frequency and

intensity of such responses which are asserted to vary as a function of the instruction stimulus, these responses become associated with the experimental situation and provide the essential cues mediating recall on the test of retention Hence, intentional learners surpass incidental learners in the amount retained, within the limits imposed by the nature of the materials and the orienting task

Given the assumed effects of the instruction stimulus on the *S*'s responses to the materials during exposure, there is nothing in the available data to suggest any systematic differences between the functional relations which characterize intentional and incidental learning This statement applies to the influence of such task variables as meaningfulness and intralist similarity and to the characteristics of recall performance Since much of the experimental work on incidental learning has used the method of free recall, the conclusion can be made more specific by asserting that the same set of principles describes free recall after intentional and incidental learning

Meaningfulness.—It has been shown that the number of items given correctly in free recall varies (*a*) with the degree of response integration, and (*b*) with the frequency of usage of integrated units (with the latter variable probably reducible to the strength of interitem associations correlated with word frequency) Increases in recall as a function of length of list are disproportionately greater when meaningfulness is high than when it is low These relationships obtain in incidental learning as well The slope of the functions relating meaningfulness and recall is, however, consistently steeper under incidental than under intentional conditions (Postman, Adams, & Phillips, 1955, Postman & Adams, 1958, Postman & Phillips, 1961) As a result, the difference between incidental and intentional learners is smaller when meaningfulness is high than when it is low This pattern of differences is found not only in the Type I studies just cited but in Type II situations as well (Bromer, 1942, Mechanic, 1962a)

An important confirmation of the generality of the relationship between meaningfulness and incidental learning comes from studies of R—S recall It is reasonable to consider R—S learning during paired associate training as a form of incidental learning of Type II since the R—S associations are not part of the task specified by the instructions (Feldman & Underwood, 1957, Jantz & Underwood, 1958) The function relating meaningfulness and R—S learning in the study of Jantz and Underwood corresponded exactly to that obtained by the method of free recall for incidental learning (Postman, Adams, & Phillips, 1955)

The interaction between meaningfulness and the conditions of learning is readily deduced from the assumption that the amount of learning depends on the frequency of effective differential responses made to the stimulus items Items of high meaningfulness are likely to evoke such responses from

intentional and incidental learners alike by virtue of the fact that they are well integrated and have high *m*-value in Noble's sense. Hence, both groups should learn such material well. As meaningfulness declines, the probability of effective differential responses falls off more rapidly under intentional than incidental conditions, and the difference in the amounts retained increases accordingly.

Sequential dependencies and serial order in recall.—It is clear that incidental learning is selective in the sense that the uninstructed *S* responds to fewer items and fewer features of the learning materials than does the instructed *S*. Such selectivity implies that the incidental learner will be less sensitive to the sequential relations between successive items than will the intentional learner. Three facts may be cited in support of this conclusion. (*a*) In immediate recall, intentional *Ss* surpass incidental *Ss* in the reconstruction of serial order even when the instructions to the former make no reference to the order of the items (Postman, Adams, & Bohm, 1956). (*b*) Increases in recall as a function of the order of approximation to the statistical structure of English are considerably smaller under incidental than under intentional conditions. There are corresponding differences in the amounts of sequential clustering during the recall of higher order passages (Postman & Adams, 1960). (*c*) Serial position curves in free recall consistently show a more pronounced finality effect under incidental than under intentional conditions (Postman & Phillips, 1954; Postman & Adams, 1957, 1960).

These findings can be subsumed under the generalization that sequential dependencies are discriminated less effectively by incidental than by intentional subjects. The failure to respond to sequential relationships is in turn reflected in the order in which the learned items are emitted in recall. Thus, in the light of the analysis of Deese and Kaufman (1957), the relative absence of a primacy effect in free recall indicates a low degree of sequential organization and an order of recall determined essentially by the sheer strength of the responses. The limited gains as a function of approximation to English again imply a low degree of sequential organization and also suggest that the learner's success in reconstructing a redundant sequence suffers when the statistical characteristics of the series have not been adequately discriminated. It should be emphasized, however, that these differences between intentional and incidental learning again are in the degree and not in the kind of functional relationships.

The selectivity of incidental learning.—Once the selectivity of the incidental *S*'s response to the stimulus materials has been taken into account, both similarities and differences between intentional and incidental learning readily fall into place without recourse to special explanatory principles. A few additional examples will be given to support this general point.

The fact of selectivity implies that the incidental learner will acquire not only fewer correct responses than the intentional learner but also fewer incorrect ones, i.e., that he will be subject to less interference from stimulus and response generalization. Consistent with this expectation, intralist similarity is less effective under incidental than under intentional conditions, and stronger remote associations are developed by the intentional learners (Postman & Adams, 1957). The fact that the "isolation effect" fails to occur under incidental conditions may likewise be attributed to a reduction in intraserial interferences (Postman & Phillips, 1954).

Increases as a function of the number of presentations occur less regularly and predictably in incidental than in intentional learning. While significant gains have been reported in some studies (G. H. Brown, 1954; Mechanic, 1962a), only small effects have been observed in other experiments (Postman & Adams, 1958; Saltzman & Atkinson, 1954). It is likely that in each case the characteristics of the materials and the conditions of the orienting task determine whether new items are responded to, and the responses to old ones are reinforced, on successive incidental trials. By contrast, such is invariably the case in instructed learning. The same conclusions appear to apply to the effects of rate of presentation (Neimark & Saltzman, 1953; Rosenberg, 1959).

Like intentional learning, incidental learning is subject to both retroactive and proactive inhibition (Gleitman & Kamrin, 1957; Postman & Adams, 1956a; Prentice, 1943; Rosenberg, 1961). The amounts of interference that are observed can be fully accounted for in terms of the degree of learning attained under the two conditions. There is also no evidence for differential rates of forgetting after intentional and incidental learning when degree of learning is taken into account (Biel & Force, 1943; Postman & Phillips, 1961).

Conclusions

Except for purposes of convenient reference to experiments in which the instruction stimulus is manipulated, there is little or no reason to maintain a conceptual distinction between intentional and incidental learning. What is learned depends on the responses elicited by the stimuli in the experimental situation. Manipulation of the instruction stimulus represents only one of the many different ways in which these responses can be determined.

Our evaluation of the conceptual status of incidental learning parallels the conclusion which we reached about the difference between short-term and long term memory. In neither area of investigation do the data seem to justify the formulation of special explanatory principles which apply to only a circumscribed range of phenomena, delimited in the one case along

the temporal dimension and in the other case by the absence of intent Both types of investigation differ in important ways from classical rote learning procedures because they are typically concerned with learning and retention under minimal conditions of reinforcement Nevertheless, they appear to be governed by the same basic principles

References

Aborn M & Rubenstein H. Information theory and immediate recall. *J exp Psychol* 1952 44, 260–266

Anderson N S Poststimulus cuing in immediate memory *J exp Psychol* 1960 60, 216–221

Averbach, E. & Coriell, A S Short term memory in vision *Bell System Tech J* 1961 60 309 328

Bahrick, H P Incidental learning under two incentive conditions *J exp Psychol* 1954 47, 170–172

Bahrick H P Fitts P M & Rankin R E Effect of incentives upon reactions to peripheral stimuli *J exp Psychol* 1952 44 400–406

Battig W F Comparison of anticipation and recall methods of paired associate learning *Psychol Rep* 1961 9, 59–65

Beebe H H Auditory memory span for meaningless syllables *J Speech Disorders* 1944 9, 273–276

Bergstrom J A Effect of changes in the time variables in memorizing together with some discussion of the technique of memory experimentation. *Amer J Psychol* 1907 18, 206–238

Hill W C & Force, R C Retention of nonsense syllables in intentional and incidental learning. *J exp Psychol* 1943 32, 52–63

Bousfield W A The occurrence of clustering in the recall of randomly arrayed associates. *J gen Psychol* 1953 49, 229–240

Bousfield W A General review of a program of research on associative clustering. *Minnesota Conference on Associative Process in Verbal Behavior* 1955

Bousfield W A., & Cohen B H The effects of reinforcement on the occurrence of clustering in the recall of randomly arrayed associates *J Psychol* 1953 36, 67 81

Bousfield W A. & Cohen B H The occurrence of clustering in the recall of randomly arrayed words of different frequencies of usage. *J gen Psychol* 1955 52 83 95

Bousfield W A. & Cohen B H Clustering in recall as a function of the number of word-categories in stimulus-word lists *J gen Psychol.*, 1956 54, 95–106

Bousfield W A. Cohen B H & Silva J G The extension of Marbe s law to the recall of stimulus-words. *Amer J Psychol* 1956 69, 429–433

Bousfield W A. Cohen B H & Whitmarsh G A. Associative clustering in the recall of words of different taxonomic frequencies of occurrence *Psychol Rep* 1958 4 39–44

Bousfield W A. Whitmarsh, G A., & Berkowitz, H Partial response identities in associative clustering. *J gen Psychol* 1960 63, 233 238

Bousfield, W A Whitmarsh G A & Esterson J Serial position effects and the "Marbe effect" in the free recall of meaningful words. *J gen Psychol.*, 1958 59, 255–262.

Brener, L R An experimental investigation of memory span *J exp Psychol*, 1940, 26, 467–482

Broadbent, D E The role of auditory localization in attention and memory span *J exp Psychol* 1954, 47, 191–196

Broadbent, D E Successive responses to simultaneous stimuli *Quart J exp Psychol* 1956, 8, 145–152

Broadbent, D E Immediate memory and simultaneous stimuli *Quart J exp Psychol*, 1957, 9, 1–11 (a)

Broadbent, D E A mechanical model for human attention and immediate memory *Psychol Rev*, 1957, 64, 205–215 (b)

Broadbent, D E. *Perception and communication* New York Pergamon Press, 1958

Broadbent, D E, & Gregory, M On the recall of stimuli presented alternately to two sense organs *Quart J exp Psychol* 1961, 13, 103–109

Bromer J A A comparison of incidental and purposeful memory for meaningful and nonsense material *Amer J Psychol*, 1942, 55, 106–108

Brown G H Factors influencing incidental learning *J exp Psychol*, 1954, 47, 163–169

Brown, J The nature of set to learn and of intra material interferences in immediate memory *Quart J exp Psychol*, 1954, 6, 141–148

Brown, J Some tests of the decay theory of immediate memory *Quart J exp Psychol* 1958, 10, 12–21

Calhoon, S W Influence of syllabic length and rate of auditory presentation on ability to reproduce disconnected word lists *J exp Psychol*, 1935, 18, 612–620 (a)

Calhoon, S W A comparison of ability to reproduce one syllable words and digits auditorially presented *J exp Psychol* 1935, 18, 621–632 (b)

Carlson, J B, & Duncan C P A study of autonomous change in the memory trace by the method of recognition *Amer J Psychol* 1955, 68, 280–284

Clark, L. L, Lansford T G, & Dallenbach K M Repetition and associative learning *Amer J Psychol* 1960, 73, 22–40

Cofer C N A study of clus ering in free recall based on synonyms *J gen Psychol*, 1959, 60, 3–10

Coleman, E B Approximations to English some comments on the method *Amer J Psychol*, 1963, 76, 239–248

Conrad, R Decay theory and immediate memory *Nature* 1957, 179, 831–832

Conrad R Errors of immediate memory *Brit J Psychol*, 1959, 50, 349–359

Conrad R Very brief delay of immediate recall *Quart J exp Psychol* 1960, 12, 45–47 (a)

Conrad, R Serial order intrusions in immediate memory *Brit J Psychol*, 1960, 51, 45–48 (b)

Conrad, R, & Hille, B A The decay theory of immediate memory *Canad J Psychol* 1958 12, 1–6

Crannell, C W, & Parrish J M A comparison of immediate memory span for digits, letters and words *J Psychol*, 1957, 44, 319–327

Crumbaugh, J C Temporal changes in the memory of visually perceived form *Amer J Psychol*, 1954, 67, 647–658

Deese, J Serial organization in the recall of disconnected items. *Psychol Rep*, 1957, 3, 577–582

Deese, J Influence of inter item associative strength upon immediate free recall *Psychol Rep* 1959, 5, 305–312. (a)

Deese, J On the prediction of occurrence of particular verbal intrusions in immediate recall. *J. exp Psychol*, 1959, 58, 17–22 (b)

Deese, J Frequency of usage and number of words in free recall: The role of association *Psychol. Rep*, 1960, 7, 337–344

Deese, J. From the isolated verbal unit to connected discourse. In C. N Cofer (Ed) *Verbal learning and verbal behavior.* New York· McGraw-Hill, 1961.

Deese, J, & Kaufman, R A. Serial effects in recall of unorganized and sequentially organized materials *J. exp. Psychol*, 1957, 54, 180–187.

Estes, W. K. Statistical theory of distributional phenomena in learning *Psychol Rev*, 1955, 62, 369–377.

Estes, W. K, Hopkins, B L, & Crothers, E J. All-or none and conservation effects in the learning and retention of paired associates *J. exp. Psychol*, 1960, 60, 329–339

Feldman, S M., & Underwood, B J Stimulus recall following paired associate learning. *J exp Psychol*, 1957, 53, 11–15.

Fraisse, P. Études sur la memoire immediate III. L'influence de la vitesse de la présentation et de la place des elements. La nature du présent psychologique. *L'Annee Psychol*, 1945, 45, 29–42

Fraisse, P., & Fraisse, R. Études sur la memoire immédiate I. L'apprehension des sons *L'Annee Psychol*, 1937, 38, 48–85

Fraser, D C. Decay theory and immediate memory *Nature,* 1958, 182, 1163.

Gates, A I, & Taylor, G A. An experimental study of the nature of improvement resulting from practice in a mental function. *J educ Psychol*, 1925, 16, 583–592

Gates, A. I. The mnemonic span for visual and auditory digits *J. exp. Psychol.*, 1916, 1, 393–403

Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning *Psychol Rev*, 1940, 47, 196–229

Gibson, J. J. The reproduction of visually perceived forms. *J exp Psychol*, 1929, 12, 1–39.

Gleitman, H, & Gillett, E The effect of intention upon learning. *J. gen Psychol*, 1957, 57, 137–149

Gleitman, H, & Kamrin, R. Proactive and retroactive inhibition in intentional and incidental learning. *Psychol Rep*, 1957, 3, 155–160

Gonzalez, R. C., & Cofer, C N Exploratory studies of verbal context by means of clustering in free recall. *J genet Psychol*, 1959, 95, 293–320

Graham, C H Behavior and the psychophysical methods An analysis of some recent experiments *Psychol Rev* 1952, 59, 62–70

Gray, J A., & Wedderburn A I Grouping strategies with simultaneous stimuli. *Quart J exp Psychol* 1960 12, 180–184

Guilford, J P., & Dallenbach, K M The determination of the memory span by the method of constant stimuli *Amer J Psychol*, 1925, 36, 621–628

Hall, J F Learning as a function of word frequency *Amer J Psychol*, 1954, 67, 138–140

Hanawalt, N G Memory trace for figures in recall and recognition. *Arch Psychol, N Y*, 1937, Whole No 216

Harden, L M The quantitative study of the similarity factor in retroactive inhibition. *J gen Psychol*, 1929, 2, 421–430

Hebb D O *The organization of behavior* New York Wiley, 1949

Hebb, D O Distinctive features of learning in the higher animal. In J F. Delafresnaye (Ed). *Brain mechanisms and learning* Oxford Blackwell, 1961.

Hebb, D O, & Foord, E N Errors of visual recognition and the nature of the trace *J exp Psychol*, 1945, 35, 335–348
Horowitz, L M Free recall and ordering of trigrams *J exp Psychol* 1961, 62, 51–57
Irwin, F. W, & Rovner, H Further study of the method of comparison applied to the problem of memory changes *J exp Psychol* 1937, 21, 533–544
Irwin, F W, & Seidenfeld, M A The application of the method of comparison to the problem of memory change *J exp Psychol*, 1937, 21, 363–381
Jantz, E M, & Underwood, B J R S learning as a function of meaningfulness and degree of S R learning *J exp Psychol*, 1958, 56, 174–179
Jenkins, J J, Mink, W D, & Russell, W A Associative clustering as a function of verbal association strength *Psychol Rep*, 1958, 4, 127–136
Jenkins, J J, & Russell, W A Associative clustering during recall *J abnorm soc Psychol*, 1952 47, 818–821
Jacobs, J Experiments in prehension" *Mind* 1887, 12, 75–79
Karlin, L, & Brennan, G Memory for visual figures by the method of identical stimuli *Amer J Psychol*, 1957, 70, 248–253
Kausler, D H, & Trapp, E P Motivation and cue utilization in intentional and incidental learning *Psychol Rev*, 1960 67, 373–379
Kausler, D H, Trapp, E P, & Brewer C L Intentional and incidental learning under high and low emotional drive level *J exp Psychol*, 1959, 58, 452–455
Kay, H, & Poulton E C Anticipation in memorizing *Brit J Psychol* 1951, 42, 34–41
Kennelly, T W The role of similarity in retroactive inhibition *Arch Psychol*, *N Y* 1941, 37, Whole No 260
Keppel, G, & Underwood B J Proactive inhibition in short term retention of single items *J verb Learn verb Behav* 1962, 1, 153–161
Koffka, K *Principles of Gestalt psychology* New York Harcourt, Brace 1935
Kohler, W, & von Restorff, H Analyse von Vorgangen im Spurenfeld II Zur Theorie der Reproduktion *Psychol Forsch* 1935 21 56 112
Lloyd, K E Short term retention as a function of average storage load *J exp Psychol*, 1961, 62, 632
Lloyd, K E, Reid, L S, & Feallock, J B Short term retention as a function of the average number of items presented *J exp Psychol* 1960, 60, 201–207
Lockhead G R Methods of presenting paired associates *J verb Learn verb Behav*, 1962, 1, 62–65
Lovibond, S H A further test of the hypothesis of autonomous memory trace change *J exp Psychol*, 1958, 55, 412–415
Lumley, F H, & Calhoon, S W Memory span for words presented auditorially *J appl Psychol*, 1934, 10, 773–784
McGeoch, J A *The psychology of human learning* New York Longmans, Green 1942
Mackworth J F Paced memorizing in a continuous task *J exp Psychol*, 1959, 58, 206–211
Mackworth N H & Mackworth J F Remembering advance cues during searching *Brit J Psychol*, 1959 50, 207–222
Marks, M R, & Jack, O Verbal context and memory span for meaningful materials *Amer J Psychol* 1952 65, 298–300
Martin, P R, & Fernberger, S W Improvement in memory span *Amer J Psychol* 1929, 41, 91–94

Mathews R Recall as a function of number of classificatory categories. *J exp Psychol* 1954 47, 241–247
Mechanic, A The distribution of recalled items in intentional and incidental learning. *J exp Psychol* 1962, 63, 593–600 (a)
Mechanic, A. Effects of orienting task, practice, and incentive on simultaneous incidental and intentional learning. *J exp Psychol*, 1962, 64, 393–399 (b)
Melton, A. W & Irwin, J McQ The influence of degree of interpolated learning on retroactive inhibition and the overt transfer of specific responses. *Amer J Psychol* 1940 53, 173–203
Melton A. W & von Lackum, W J Retroactive and proactive inhibition in retention Evidence for a two-factor theory of retroactive inhibition *Amer J Psychol* 1941, 54, 157–173
Metraux, R. W Auditory memory span for speech sounds *J Speech Disorders* 1944 9, 31–38
Miller, G A The magical number seven plus or minus two Some limits on our capacity for processing information *Psychol Rev* 1956, 63, 81–97
Miller, G A Free recall of redundant strings of letters *J exp Psychol*, 1958, 56, 485–491
Miller, G A & Selfridge, J A. Verbal context and the recall of meaningful material *Amer J Psychol* 1950 63, 176–185
Moray N Broadbent s filter theory Postulate H and the problem of switching time *Quart J exp Psychol*, 1960, 12, 214–220
Mulhall, E F Experimental studies in recall and recognition. *Amer J Psychol* 1915 26, 217–228
Muller, G E., & Pilzecker, A. Experimentelle Beitrage zur Lehre vom Gedachtniss. *Z Psychol* 1900 Erbd. I 1–300
Murdock, B B Jr. The immediate retention of unrelated words. *J exp Psychol.*, 1960 64, 222–234
Murdock, B B Jr. The retention of individual items *J exp Psychol*, 1961 62, 618–625
Myers, G C. A study in incidental memory *Arch Psychol N Y* 1913, 4, Whole No 26
Neimark, E. & Saltzman I J Intentional and incidental learning with different rates of stimulus presentation *Amer J Psychol.*, 1953 66, 618–621
Newton J M., & Wickens D D Retroactive inhibition as a function of the temporal position of interpolated learning. *J exp Psychol* 1956, 51, 149–154
Oberly H S. A comparison of the spans of "attention" and memory *Amer J Psychol* 1928 40, 295–302
Peatman J G & Locke, N M Studies in the methodology of the digit span test. *Arch Psychol N Y* 1934 25, Whole No 167
Peterson L. R. Immediate memory Data and theory In C. N Cofer & B S. Musgrave (Eds.) *Verbal behavior and learning Problems and processes* New York McGraw Hill 1963 Pp 336–353
Peterson L. R., & Peterson M J Short term retention of individual verbal items. *J exp Psychol.*, 1959 58, 193–198
Peterson L R & Peterson M J The effect of spacing repetitions on short term retention. *Amer J Psychol* 1960 15, 450 (abstract)
Peterson L R Peterson M J & Miller, A Short term retention and meaningfulness. *Canad J Psychol* 1961 15, 143–147
Peterson L R., Saltzman D Hillner, K & Land V Recency and frequency in paired associate learning. *J exp Psychol* 1962, 63, 396–403

PILLSBURY, W B, & SYLVESTER, A Retroactive and proactive inhibition in immediate memory *J exp Psychol*, 1940, 27, 532–545

POLLACK, I Assimilation of sequentially encoded information *Amer J Psychol*, 1953, 66, 421–435

POLLACK, I, JOHNSON, L B, & KNAFF, P R Running memory span *J exp Psychol*, 1959, 57, 137–146

PORTER, E H, JR The influences of delayed instructions to learn upon human performance *J exp Psychol*, 1938, 23, 633–640

POSTMAN, L Repetition and paired associate learning *Amer J Psychol*, 1962, 75, 372–389 (a)

POSTMAN, L The temporal course of proactive inhibition for serial lists *J exp Psychol* 1962, 63, 361–369 (b)

POSTMAN, L Transfer of training as a function of experimental paradigm and degree of first list learning *J verb Learn verb Behav*, 1962, 1, 109–118 (c)

POSTMAN, L, & ADAMS P A Studies in incidental learning III Interserial interference *J exp Psychol* 1956, 51, 323–328 (a)

POSTMAN, L, & ADAMS, P A Studies in incidental learning IV The interaction of orienting tasks and stimulus materials *J exp Psychol*, 1956, 51, 329–333 (b)

POSTMAN, L, & ADAMS, P A Studies in incidental learning VI Intraserial interference *J exp Psychol*, 1957, 54, 153–167

POSTMAN, L, & ADAMS, P A Studies in incidental learning VII Effects of frequency of exercise and length of list *J exp Psychol* 1958, 56, 86–94

POSTMAN, L, & ADAMS, P A Studies in incidental learning VIII The effects of contextual determination *J exp Psychol* 1960 59, 153–164

POSTMAN, L, ADAMS P A, & BOHM, A M Studies in incidental learning V Recall for order and associative clustering *J exp Psychol*, 1956, 51, 334–342

POSTMAN L, ADAMS P A, & PHILLIPS, L W Studies in incidental learning II The effects of association value and of the method of testing *J exp Psychol*, 1955, 49, 1–10

POSTMAN, L, & PAGE, R Retroactive inhibition and psychophysical judgment *Amer J Psychol* 1947, 60, 367–377

POSTMAN, L, & PHILLIPS, L W Studies in incidental learning I The effects of crowding and isolation *J exp Psychol*, 1954, 48, 48–56

POSTMAN, L & PHILLIPS, L W Studies in incidental learning IX A comparison of the methods of successive and single recalls *J exp Psychol* 1961, 61, 236–241

POSTMAN, L, & RILEY, D A Degree of learning and interserial interference in retention *Univer Calif Publ Psychol* 1959, 8, 271–396

POSTMAN, L, & SENDERS, V L Incidental learning and generality of set *J exp Psychol* 1946, 36, 153–165

POULTON, E C Memorization during recall *Brit J Psychol*, 1953, 44, 173–176

PRATT, C C The time-error in psychophysical judgments *Amer J Psychol*, 1933, 45, 292–297

PRATT, C C The law of disuse *Psychol Rev*, 1936, 43, 83 93

PRENTICE, W C H Retroactive inhibition and the motivation of learning *Amer J Psychol*, 1943, 56, 283–292

RAFFEL, G Two determinants of the effects of primacy *Amer J Psychol* 1936, 48, 654–657

REID L S, LLOYD, K E, BRACKETT, H R, & HAWKINS, W F Short term retention as a function of average storage load and average load reduction *J exp Psychol* 1961, 62, 518–522

RILEY D A Memory for form In L Postman (Ed), *Psychology in the making* New York Knopf, 1962

ROBINSON, E S The "similarity factor in retroaction *Amer J Psychol*, 1927, 39, 297–312

ROCK, I The role of repetition in associative learning *Amer J Psychol*, 1957, 70, 186–193

ROHRER, J H Factors influencing the occurrence of reminiscence Attempted formal rehearsal during the interpolated period *J exp Psychol*, 1949, 39, 484–491

ROSENBERG, S Exposure interval in incidental learning *Psychol Rep*, 1959, 5, 675

ROSENBERG, S Retroactive inhibition in incidental learning. *Amer J Psychol*, 1961, 74, 283–286

ROTHKOPF, E Z., & COKE, E U The prediction of free recall from word association measures *J exp Psychol*, 1961, 62, 433–438

RUBENSTEIN, H, & ABORN, M Immediate recall as a function of degree of organization and length of study period *J exp Psychol*, 1954, 48, 146–152

RUSSELL, W A, & JENKINS, J J Associative clustering and response rank in the Kent Rosanoff test. *ONR Tech Rep*, No 2, University of Minnesota, 1952.

SALTZMAN, I J The orienting task in incidental and intentional learning. *Amer J Psychol* 1953, 66, 593–597

SALTZMAN, I J Comparisons of incidental and intentional learning with different orienting tasks *Amer J Psychol*, 1956, 69, 274–277

SALTZMAN, I J, & ATKINSON, R. L Comparison of incidental and intentional learning after different numbers of stimulus presentations *Amer J Psychol*, 1954, 67, 521–524

SHARP, H C Effect of contextual constraint upon recall of verbal passages *Amer J Psychol* 1958 71, 568–572

SHELLOW, S M Individual differences in incidental memory *Arch Psychol N Y*, 1923, 10, Whole No 64

SHEPARD, R. N, & TEGHTSOONIAN, M Retention of information under conditions approaching a steady state *J exp Psychol*, 1961, 62, 202–222

SILVERMAN, R. E Anxiety and the mode of response *J abnorm soc Psychol*, 1954, 49, 538–542

SILVERMAN, R. E & BLITZ, B Learning and two kinds of anxiety *J abnorm soc Psychol* 1956, 52, 301–303

SPERLING, G The information available in brief visual presentations. *Psychol Monogr*, 1960 74, Whole No 498

TWEDT, H. M, & UNDERWOOD, B J Mixed vs. unmixed lists in transfer studies *J exp Psychol* 1959, 58, 111–116

UNDERWOOD B J The effect of successive interpolations on retroactive and proactive inhibition *Psychol Monogr* 1945 59, Whole No 273

UNDERWOOD B J Proactive inhibition as a function of time and degree of prior learning *J exp Psychol* 1949 39, 24–34

UNDERWOOD B J Stimulus selection in verbal learning. In C N Cofer & B S Musgrave (Eds.) *Verbal behavior and learning Problems and processes* New York McGraw Hill 1963 Pp 33–48

UNDERWOOD B J Degree of learning and the measurement of forgetting. *J verb Learn verb Behav* in press.

UNDERWOOD B J., RUNQUIST W N & KEPPEL G Item selection in paired associate learning. *Amer J Psychol* 1962 75, 353–371

UNDERWOOD B J., RUNQUIST W N & SCHULZ, R. W Response learning in paired

associate lists as a function of intralist similarity *J exp Psychol*, 1959, **58**, 70–78
UNDERWOOD, B J, & SCHULZ, R W *Meaningfulness and verbal learning* Philadelphia Lippincott, 1960
WAUGH, N C Serial position and the memory span *Amer J Psychol*, 1960, **73**, 68–79
WAUGH, N C Free versus serial recall *J exp Psychol*, 1961, **62**, 496–502
WELCH, G B, & BURNETT, C T Is primacy a factor in association formation? *Amer J Psychol*, 1924, **35**, 396–401
WOODWORTH, R S *Experimental psychology* New York Holt 1938
WULF, F Uber die Veranderung von Vorstellungen *Psychol Forsch*, 1922, **1**, 333–373
YNTEMA, D B, & MUESER, G E Remembering the present state of a number of variables *J exp Psychol* 1960, **60**, 18–22
YNTEMA, D B, & MUESER, G E Keeping track of variables that have few or many states *J exp Psychol*, 1962, **63**, 391–395
YOUNG, C W, & SUPA, M Mnemonic inhibition as a factor in the limitation of the memory span *Amer J Psychol* 1941, **54**, 546–552
ZANGWILL, O L An investigation of the relationship between the processes of reproducing and recognizing simple figures *Brit J Psychol*, 1937, **27**, 250–276

Behavioral Effects of Instruction to Learn:

COMMENTS ON PROFESSOR POSTMAN'S PAPER

JAMES DEESE
The Johns Hopkins University

Postman's review and analysis of the two topics of short-term memory and incidental learning are remarkably complete and detailed Indeed, a first reaction upon reading the chapter is that there is little more worth saying One can find only a few minor issues about which to quarrel, and it is difficult to find anything to which to object in the general conclusions reached by Postman Therefore, the major portion of these comments is directed toward bringing the work reviewed by Postman together in a slightly different way While such comments probably serve to add little to the substance of Postman's chapter, it may help to bring out more strongly some of the implications of his analysis

Postman approaches incidental learning through short term memory I shall reverse the order and approach the problems of short-term memory by way of the questions raised in the study of incidental learning Although Postman finds empirical grounds for relating these topics, he does not, I think, make clear one implication of his analysis of incidental learning for problems in the study of short term memory

INCIDENTAL LEARNING

The basic problem in the analysis of incidental learning is provided by the nature of the differences in the things people do under instructions to learn compared with no instructions to learn when exposed to the same materials Postman summarizes the basic differences very well when he says "Analysis of the effects of the orienting task has led us to the conclusion that intent per se is not a significant variable in learning [Indeed, one may ask, does the concept of intent per se have any psychological meaning at all?] but that the instruction stimulus influences the amount and characteristics of learning by determining the differential cue-producing responses including deliberate rehearsal, which occur during the period of practice These differential responses include categorizing responses such as naming or labeling, other responses elicited by the items

through stimulus generalization, and responses serving as associative links among the members of a series It is the frequency and intensity of such responses which are asserted to vary as a function of the instruction stimulus" (pp 190–191)

Here Postman states the basis of the difference between *incidental* and *intentional* learning and describes the conditions responsible for the obtained experimental differences produced by varying the instructions to learning I think, however, we may simplify Postman s description of the differential responses aroused during learning in order to state those that are most essential There is a class of responses which is basic, not so much because it is important in describing the experimentally obtained differences, but because, in theory, it defines the occasion for learning and thus provides opportunity for the alteration in the behavior of a subject when he is exposed to material after particular instructions

Under both incidental and intentional conditions *S*s are exposed to a set of materials In most of the studies described by Postman, the materials are quite familiar to the *S* (words, letters, numbers, etc) The task of *S* is to emit these materials in some prescribed order in the absence of direct cues

The experimenter insures, to the best of his ability to design an experiment, that *S*s under both incidental and intentional conditions perceive the material exposed to them That is to say, the experiment produces a set of conditions which almost certainly demand that *S*s react to the critically discriminable aspects of the material presented If he does not, the comparison between incidental and intentional conditions is contaminated by a difference in the amount of information taken in by the *S*s Even when the experimental comparisons are carefully controlled, however, as the literature shows with almost complete if not with complete unanimity, *S*s do not recall material as well under incidental as under intentional instructions

Under intentional conditions *S*s are instructed, in one way or another, to learn The major problem in the analysis of incidental learning is not in the conditions of incidental learning per se but in determining what alterations in behavior are induced by the formal instructions to learn something It is to this problem that I wish to address an hypothesis

The hypothesis to be stated here is a rather naive one, it is simple enough to be almost certainly wrong, if it is taken as a statement of the only condition produced by instructions to learn It may, however, be the most important condition and the one that is critical to the differences between intentional and incidental learning

The hypothesis may be stated as follows Instructions to learn (or to memorize) produce a set to emit "representational" responses to the

material presented These representational responses are responsible for the encoding of the presented material The term representational response I have borrowed from Bousfield He uses it to describe responses made to verbal materials that are identical in linguistic form to the materials themselves The concept has been used by Bousfield, by Cofer, and by others in the description of certain relations among free associations

Bousfield has limited the representational response to verbal stimuli and responses, and he has assumed that it is always identical, in form, to the presented stimulus As such, however, the representational response is a special case, and the general case is any identifying response that serves to encode presented material, whether that material is verbal or not

The identifying code for a linguistic stimulus is, by definition, identical in form to that stimulus The code for a picture or an abstract form, of course, is not Nevertheless, a verbal code for such a stimulus may serve as a device for remembering the salient features of the stimulus, and, if needed, it may provide for the reconstruction of the general appearance of the stimulus Rather than imply, however, that every identifying response is a *name* for something, I would prefer Bousfield s neutral and descriptive term, representational response

Representational responses may be either overt or implicit, if they are implicit they are unarticulated phonemic equivalences of some potentially overt response, and, as such, bear some simple direct relation to visually or aurally presented material They are not merely visual or auditory stimulus traces, however, a point that is important in considering some of the problems of short-term memory

Instructions to learn, then, produce appropriate representational responses during presentation of material, conditions of incidental learning must also produce some such responses, but these are fewer in number The greater frequency of representational responses during intentional learning provides one, perhaps the essential, difference between performance under incidental learning and intentional learning Another is provided by the rehearsal of responses during the absence of stimulus presentation, little or no such rehearsal is likely to occur for conditions of incidental learning, even when appropriate representational responses are made during presentation

The hypothesis of representational responding leads to some apparent difficulties with the experimental literature There are a few (and to my knowledge only a few) experiments in incidental learning in which subjects are forced to make *overt* identifying responses under conditions of incidental learning These studies (Jenkins, 1933, Brown, 1954, Postman

& Phillips, 1954) uniformly show the expected *superiority of intentional* conditions

I would like to be able to say that these studies uniformly show a very high level of performance under incidental conditions, even though it is not quite so high as that under intentional conditions So much would be consistent with the supposed importance of representational responding, since intentional learning would produce some implicit responding (re hearsal) not present under incidental conditions I would also suspect that what differences there were between incidental and intentional learning would be greatest after earlier stages of practice or presentation Such a result would be a consequence of supposing that learning the members of the respones set, irrespective of order or position, would place the greatest advantage upon differences in frequencies of emission of representational responding

In a similar vein, an experiment modelled after that of Postman and Adams (1960) should produce the smallest possible difference between incidental and intentional conditions In the first of the two experiments described by these authors, material was presented to *S*s for immediate free recall The material consisted of selected orders of approximation to English from the Miller-Selfridge lists The *S*s under incidental instructions were given a rating cover task, while the intentional *S*s were instructed to remember the material As expected, there was a difference for all orders of approximation The difference was a function of orders of ap proximation and roughly a constant proportion of total recall

Suppose, however, that this experiment had been conducted on the model of the Jenkins (1933) experiment Suppose that the cover task for the incidental learners required the emission of overt representational responses at the same rate and under the same conditions (save for the instructions to remember) as for the intentional learners If so, the pattern and extent of representational responses should be very nearly the same under the two conditions There would be no anticipatory set induced by the typical role learning technique, and if the conditions of pacing were right there would be little or no opportunity for differences in rehearsal Therefore, the representational responding hypothesis would lead one to suppose that there would be very nearly equal recall for intentional and incidental instructions

If provision is made for more or less equal representational responding, there should be no difference in the ability of *S*s to use pre-existing associative linkages There is strong evidence, in some of the literature reviewed by Postman in his analysis of short term memory, that deliberate instructions to use these pre existing linkages does not increase the extent to which they are actually used by *S*s

The review presented by Postman makes almost no mention of any possible direct effects of motivation in the differences between the conditions of incidental and intentional learning. Such an omission is the result of the emphasis he places on the nature of the responses elicited by the instructions given to *Ss*. Furthermore, the empirical evidence shows very little in the way of effects which can be said to be the direct result of changes in incentive conditions. Postman's analysis of the experiments by Bahrick (1954) and Mechanic (1962) shows what effects exist to be the correlates of the pattern of responding introduced by differences in instruction. Such a view can very easily be extended to the experiments by Johnson and Thomson (1962).

Postman's conclusion leads to the implication, supported by the analysis of many problems in human learning, including both incidental learning and short-term memory, that variables are effective in controlling the amount and rate of learning to the extent that they control the distribution and frequency of appropriate responses to the stimulus material. Therefore, as Postman asserts in his conclusion, we do not need to appeal to special principles of intent or motivation or incentive value to describe the differences between intentional and incidental learning. Learning will take place to the extent that it is possible to induce *Ss* to emit the appropriate responses. There is undoubtedly an implication here for those who design programs for machine teaching.

The Stimulus Trace and Short-Term Memory

Before we turn to the role of the representational response in short-term memory, a word needs to be said about the concept of the stimulus trace. I detect, in Postman's chapter, a sense of discomfort with the idea of the stimulus trace. It is very easy to share Postman's uneasiness, for the concept leads to hopelessly difficult experimental problems. Postman's summary of the experimental literature shows that it is extremely difficult to find evidence for the existence of a "pure" stimulus trace that—to use the phrase from Gestalt theory—has autochthonous properties. Yet, the recent experiments by Sperling (1960) and Averbach and Coriell (1961) present very strong evidence for sensory storage over very brief intervals of time. The sensory storage system provides opportunity to hold perceived events until "processed" (presumably responding, overtly or implicitly), but these traces are not wiped out by the events in processing. Instead, the data from the Averbach and Coriell experiments suggest that they are wiped out by succeeding input into the same sensory channel.

Here is an opportunity for some discontinuity between the traditional conditions of long term retention and the data on short term retention

If the interval is short enough, traces may be wiped out, not by competing response systems, but by additional input of information These events, if we are to judge from the available data, are very short, but it is at least conceivable that they could extend over several seconds, which is the order of duration of some of the events studied in short term memory

Nevertheless, Postman's analysis makes it abundantly evident that the familiar effects in response interference must play an important role in short term memory, even for single presented items The extent to which such short term memory is also under the influence of autochthonous trace processes is still an open question

Beyond very short stimulus trace effects, however, short term memory must, in part at least, depend upon representational responding Such responding, however, may, in most instances, also produce some strength for items which are available for proactive and retroactive interference Proactive interference will occur when there are no cue conditions which differentiate *occasions* for the emission of one or another response In most, if not all short term memory experiments, such differential cue conditions are minimal, since contextual effects from item to item are likely to be similar, and usually there is high interitem generalization Therefore, a very small amount of material may provide a source for proactive interference To such proactive interference may be added extra experimental sources of interference, though these may be less strong in short term memory experiments because of strong contextual boundaries

When redundancy is present, of course, recall may produce items which were not responded to during presentation Depending upon the amount and distribution of the redundancy, these items may greatly enhance obtained recall It is such an effect which is responsible for the influence of interitem associations, category membership and sequential dependencies upon the measures of recall, and, because such conditions do not depend upon emitting what was responded to during presentation but *emitting responses correlated with these*, Asch and Ebenholtz (1962) do not consider such effects to be influences upon memory Nevertheless, if the constant memory span or "chunking" hypothesis is correct (and this seems to be a more attractive notion all the time, cf Waugh, 1962), such effects must be responsible for nearly all, if not all, variations in amount recalled produced by variations in presented material

In any event, Postman's review shows convincing evidence that the retention of material, particularly verbal material, over short intervals of time is strongly under the influence of the amount and nature of responses available to an individual at that time Both interference effects and enhancement of recall may be seen to be the result of the same process, responses which are either chained to responses learned by the *S* or

elicited by the contextual stimuli of the memory testing situation. Whether these effects are superimposed upon a changing trace or not is, as Postman points out, almost impossible to say from present techniques (the promise provided for very short traces by the experiments of Sperling and of Averbach and Coriell aside). The importance of these responses both for forgetting and for enhancement of recall, as in interitem association, brings the variables of short term memory squarely face to face with those of the incidental learning experiments. The major effects in both kinds of experiments are the result of the distribution of responses available to the *S* at the time of testing. The experiments differ chiefly in the way in which the availability of responses is systematically varied.

The two types of experiments also raise interesting questions about the nature of any supposed stimulus trace. To the extent that some encoding or something like the representational response is not absolutely necessary to immediate recall, the differences between incidental and intentional conditions should be reduced. Indeed, these differences should almost disappear in a recognition test of memory, if recognition memory can be based upon a pure stimulus trace.

The question of recognition memory is difficult, for the study of recognition is full of methodological pitfalls. The important study of Shepard and Teghtsoonian (1961) shows remarkable levels of correct recognition following the presentation of very long intervening series. Recognition is much more sensitive than recall for both short- and long term memory, a fact which, perhaps, may encourage those convinced of the reality of a stimulus trace. If we suppose that there is an absence of the pure effect of intent or motivation in incidental learning, the only possible source of differences between incidental and intentional learning is in the encoding of information (or the representational response, to be more specific). Therefore, we should expect that the stimulus trace under incidental conditions, particularly in Postman's Type II experiment, ought not to produce results very different from what it would under intentional conditions. In a carefully and properly designed experimental comparison of incidental and intentional learning one ought to minimize the differences between these conditions by the use of a recognition test of memory. Such a result would make more encouraging the search for the effects of a pure stimulus trace in both short- and long term memory experiments.

References

Asch, S. E., & Ebenholtz, S. M. The process of free recall: evidence for non associative factors in acquisition and retention. *J. Psychol.*, 1962, 54, 3–32.

Averbach, E., & Coriell, A. S. Short term memory in vision. *Bell System Tech. J.*, 1961, 40, 309–328.

BAHRICK, H P Incidental learning under two incentive conditions *J exp Psychol* 1954, 47, 170–172

BROWN, G H Factors influencing incidental learning *J exp Psychol*, 1954, 47, 163–169

JENKINS, J G Instruction as a factor in "incidental" learning *Amer J Psychol* 1933, 45, 471–477

JOHNSON, R, & THOMSON, C Incidental and intentional learning under three conditions of motivation *Amer J Psychol*, 1962, 75, 284–288

MECHANIC A Effects of orienting task, practice, and incentive on simultaneous incidental and intentional learning *J exp Psychol* 1962, 64, 393–399

POSTMAN, L, & ADAMS, P A Studies in incidental learning VIII The effects of contextual determination *J exp Psychol*, 1960 59, 153–164

POSTMAN, L, & PHILLIPS, L W Studies in incidental learning I The effects of crowding and isolation *J exp Psychol*, 1954, 48, 48–56

SHEPARD, R N, & TEGHTSOONIAN, M Retention of information under conditions approaching a steady state *J exp Psychol*, 1961, 62, 202–222

SPERLING, G The information available in brief visual presentations *Psychol Monogr*, 1960 74, Whole No 498

WAUGH, N C Length of series and the learning curve *Amer J Psychol* 1962 75, 177–192

The Concept of the Concept[1]

HOWARD H KENDLER *
New York University

The writer could not help but speculate, after receiving the invitation to participate in this symposium, about whether his assigned topic would have been included if everyday language were without a word such as *concept* Being so common and possessing such apparent psychological significance forces the psychologist to treat this topic as a fundamental one

This observation about the linguistic import of the term concept emphasizes the point that the present subject, as distinguished from the others in this symposium, with the exception of problem solving, is not a technical term having its origin in a clear cut experimental methodology Of course, many respectable scientific concepts do have their roots in common parlance They, however, achieve respectability and importance only after these original roots have withered and their place and function have been taken over by a technical term, or terms, that have the advantages not only of being less ambiguous but also, in the experimental sense, more meaningful Although there are signs that this healthy course of development is beginning for the concept of the concept, the fact is that it is still vague and amorphous This point is made to impress the reader that because of the scientific infancy of his topic, the writer is in that unenviable but not uncommon position of a psychologist who is not quite sure about what he is writing This predicament nevertheless does have one advantage It frees one from being constrained by any orderly array of facts and theories that demand a particular kind of systematic treatment It allows for flights of fancy and broad generalizations without much fear of being embarrassed by any clear-cut contradictory evidence

With this prologue concluded the general approach that will be made in the analysis of conceptual behavior can now be described Initially the methods used to investigate concept learning will be examined, particularly in terms of how the empirical attacks upon this problem have been

[1] The writer wishes to express his great appreciation to the Office of Naval Research and to the National Science Foundation for their support of his research efforts in the field of classificatory and problem solving behavior Thanks are also due to Tracy S Kendler for her helpful reading of the manuscript.

* Present Address *University of California Santa Barbara*

shaped by theoretical preconceptions An attempt will be made to tease out significant characteristics of these approaches so that some overview, although no doubt an incomplete one, will be offered of the current scene of research and theory in the field of classificatory behavior. Finally, the problems of conceptual behavior will be systematized in terms of stimulus-response language For the most part this analysis will be concerned not with criticism but with clarification, not with truth but with meaning

Models of Concepts

Typically, models of behavior have been extended *to* concept learning As a result there is often a greater interest in demonstrating how concept learning is like something else than it is like itself There is, of course, nothing wrong in approaching a research area with certain preconceptions, or what may be more appropriately called a model (Lachman, 1960) Such biases—and they are biases—sometimes lead to important theoretical generalizations (not to mention the comfort they can provide during lengthy periods of ignorance) But until the models pay off, it should be remembered that they are neither demanded by, nor an outgrowth of, the data These preconceptions are brought in from the *outside* Researchers must guard against the tendency to be overconcerned with demonstrating the appropriateness of the model at the expense of coping with problems, both empirical and theoretical, that characterize the behavior that is being investigated

The possible prejudicial effects of pretheoretical models are sometimes minimized because of the prevailing conviction that they can always be disentangled from empirical results Although the writer has no desire to enter the debate about whether ideally scientific language can be divided nicely and neatly into a theoretical and empirical component (Hanson, 1958), he does wish to emphasize that it is often far more difficult than scientists like to believe Theories and models often persist simply because they are impregnated with facts As long as the facts stand, the theory and model survive These comments, it should be understood, are not directed at attacking the *use* of pretheoretical models, instead a warning is being issued against their possible *misuse*

S—R Conceptions of Concepts

The initial sortie into the realm of models of concept learning will be in the direction of S—R conceptions There are two reasons for this choice The first is that S—R conceptions have probably influenced more workers in the field of concept learning than any other view Second, they serve as a frame of reference to explicate and analyze other formulations

S—R models emerge from three different orienting attitudes (1) Behavior can be represented in terms of S—R associations, (2) the science of behavior is investigated best with a sophisticated behavioristic methodological orientation, and (3) S—R theories of behavior of simple phenomena such as conditioning and discrimination learning are fruitful sources of models for more complex behavior

While no detailed analysis of these three components will be offered at present, it will be helpful to refer to certain points, particularly in relation to the meaning of stimulus-response associationism as a language system *representing* psychological events

The essence of stimulus-response language is contained in the S—R paradigm Accordingly, there are three important characteristics of behavioral events These are the stimulus or stimuli, the response or responses, and the association (the—) between them On the simplest level possible this means that these terribly complex phenomena which we call *behavior* can be represented advantageously in language of stimulus response associationism Behavioral events can be described by specifying some features of the environment (stimulus) and some component of the total behavior pattern (response), and the relationship between the two

Much of the confusion surrounding stimulus response language has resulted from overburdening the descriptive linguistic system with theoretical properties Stimulus response language need not be committed to any physiological conception of behavior, nor need it be committed to any view of how a functional relationship (Robinson, 1932) between a stimulus-response association is established or strengthened It simply is a way —and not the only way—of describing behavior

When this S—R language system is combined with a behavioristic methodological orientation, which can be succinctly characterized as physicalistic, operational, and experimental (Estes, 1959), it is not surprising that early S—R psychologists clutched to their bosoms conditioning procedures and phenomena The consequences of this action should not be ignored Not only was it decided to investigate and theorize about conditioning, but also to perceive other behavioral phenomena as if they were in some way (Hilgard & Marquis, 1940) like conditioning

The first and perhaps the most important extension of the conditioning model was to the realm of discrimination learning Pavlov (1927), in his investigations of differentiation, indicated how this phenomenon could be explained as resulting from the combined operation of excitation and internal inhibition But more important in shaping the history of the psychology of learning was Spence's (1936) classical theory of discrimination learning Conceptually it was an extension of the view that conditioning processes (conditioning, geocralization, and extinction) were operating in

discrimination learning It, in turn, extended this conditioning generalization-extinction model to new empirical realms It may not appear much of a leap from Pavlov's differentiation to Spence's discrimination Nevertheless, it proved sizable, as witnessed by the numerous problems and controversies it generated

Perhaps the major effect of taking this apparently small step from conditioning to discrimination learning was the lessening of the control by the experimenter over the subject's (*S*'s) reception of the critical stimulus In classical conditioning the *S*'s behavior is so restricted and the environment is so unchanging that the presentation of the conditioned (positive) or differentiated (negative) stimulus practically guarantees stimulus reception In contrast, in the typical discrimination problem conventionally used with rats, the *S* is moving constantly about and the critical stimulus does not occupy such a perceptually dominant position in the *S*'s environment. In addition there are two critical stimuli to which the organism must respond This usually complicates the stimulus reception process by creating special problems of stimulus patterning and receptor orienting responses

The differences between conditioned differentiation and discrimination learning can be reduced by arranging the discrimination apparatus to minimize problems of stimulus reception (e g , use black and white alleys in which no matter what the animal does, except closing his eyes, he will see black or white) The use of a successive discrimination problem instead of a simultaneous one also has the effect of making the discrimination problem more similar to the differentiation one

It can be argued that it is neither necessary nor desirable to try to investigate discrimination learning in an experimental situation arranged so that it resembles the classically conditioned differentiation If, however, one is particularly interested in extending the conditioning generalization-extinction model, there are obvious strategic advantages in initially selecting discrimination problems that appear similar to classical conditioning

The relationship between conditioning and discrimination learning, however, was not perceived only in terms of research strategy It was also considered in terms of basic theoretical issues With the help of sharp pens and thin skins theoretical discussions produced arguments and acrimony, disagreements and distortions confusions and confabulations An Olympian sage looking at these controversies in the tolerant light of history, might conclude they revolved about two related problem areas (1) whether the principles of conditioning generalization, and extinction could support a general theory of discrimination learning, and (2) whether independent empirical stimulus variables could potentially be coordinated

to the independent theoretical stimulus variable of any S—R theory (Koch, 1954)

In analyzing the first problem it is helpful to view the conditioning-generalization extinction model in both a weak and strong form In a weak form it is very little more than a simple description of discrimination phenomena expressed in terms of S—R associations A successful discrimination between two features of the environment results from learning to respond to the positive one and eliminating the generalized tendency to respond to the negative one The strong form of S—R discrimination theories specifies important additional properties the nature of the stimuli and responses as well as the associative process, the reinforcing mechanisms that form and strengthen associations, the principles of stimulus generalization, the manner in which one habit gains ascendancy over another, etc The weak and strong forms of the theory can be considered as end points on a dimension of theoretical development Quite obviously contemporary S—R discrimination theories are past the weak point but equally obviously, they have yet to reach the strong point If this appraisal is correct, then one cannot as yet decide whether or not the conditioning generalization-extinction model has the potential to develop into an adequate general theory of discrimination The proof of the pudding is in the eating and the pudding is still in the oven In order to prevent this remark from being considered as an attempt to shield S—R discrimination theories from possible criticism, it may be appropriate to state flatly that S—R discrimination theories are much weaker than they are made to appear A visit to the laboratory with an intention to test S—R discrimination theories will reveal that fairly exact predictions can be made in only the simplest of situations

The second problem, which is that of defining a stimulus, is one that is constantly with us The view that it is insoluble, or that it lends itself to only one possible solution seems premature At the same time no amount of turning one's back on this problem will make it disappear (This symposium, with its frequent reference to the problem of the stimulus, testifies to its existence) In classical conditioning and operant conditioning with discrimination stimuli, neat relationships are obtained between a physically defined stimulus and behavior No definitional problem seems to exist Even when a simple physical definition of the stimulus fails to produce a tidy empirical relationship, it is possible to cope with such a problem, as did Hovland (1937), by redefining the stimulus dimension in terms of different mathematical properties Although such derived measures can assume many forms requiring both ingenuity and mathematical sophistication, epistemologically they share the common property of defining a stimulus as a physical event Looking at the problem of stimulus definition

within the limited scope of conditioning phenomena, it is easy to see how it is possible to conclude mistakenly that the problem is nonexistent

When one leaves the security of conditioning and enters the uncertainty of discrimination learning, it becomes apparent that some special mechanisms are required to select out the appropriate stimulus Such concepts as receptor orienting or observing responses, attention, stimulus patterning, and response produced cues all function as selective associative mechanisms

In experimental situations in which the relevant stimuli do not "surround' the *S*, he must first learn to make appropriate receptor orienting acts (e g , Ehrenfreund, 1948) The problem of attention is brought home in the physiological work (Hernandez Peon, Scherrer, & Jouvet, 1956) showing that a distinctive and obvious sound can be prevented from eliciting neural impulses in the auditory nerve when the organism is attending to something else Since Pavlov (1927) it has been recognized that all elements of a stimulus compound are not equally important in associative formation Whether an explanation of stimulus patterning can be deduced from fundamental conditioning principles as Hull (1943) had hoped, or whether it depends on independent perceptual principles as Lashley (1938) and Krechevsky (1938) indicated, or whether a learning perceptual orientation can be combined as Dodwell (1961) suggests, is at present a moot question But it is quite clear that the stimulus patterning problem must be solved if the S—R discrimination theory is to become stronger"

The final line of evidence that focuses attention upon the problem of defining a stimulus is the one that will concern us most It has to do with defining the stimulus situation in terms of the cues resulting from responses, both explicit and implicit, of the organism There has been a definite trend in many S—R theories (e g , Spence 1956) for large parts of the environment to which the S is responding to be introjected into the organism Historically this trend bears some relationship to Gestalt psychologists' distinction (e g , Koffka 1935) between the behavioral (psychological) and geographical (physical) environments a distinction that was offered to illustrate the right and wrong way to proceed in analyzing behavior in relation to the environment Bergmann neatly gets to the heart of the matter 'But even so what is the predictive value of the suggestive metaphor psychological environment ? Is it not the business of science to ascertain which objective factors in the past and the present states of the organism and its environment account for the difference in response, so that we can actually predict it instead of attributing it, merely descriptively and after it has happened, to a difference in the psychological environment?" (Bergmann 1943 p 133)

This sort of criticism, it should be realized, is not basically critical of any construction of an environment that intervenes between the physical environment and overt behavior The weight of the criticism is directed at ad hoc constructions Postulating some "intervening" environment is permissible and potentially useful if (1) its conceptual properties can be described in an a priori manner, and (2) its influence on behavior can be specified These are the goals that S—R formulations have sought to reach when postulating mechanisms like the anticipatory goal response (Hull, 1930) which in effect shifts attention from an external to an internal environment These hypothetical intervening environments have resulted in part from the prodding of, and reaction to, cognitive (Tolman, 1932) and Gestalt-type formulations of an internal environment, even though it should be recognized, as Goss has shown (1961), that the earliest behaviorists (e g , Max Meyer, John B Watson) assumed an intervening environment in the form of mediating verbal responses

The analysis of the expansion of the stimulus construct serves to lay the groundwork for beginning our analysis of conceptual behavior You recall that the extension of conditioning theory to discrimination behavior, an empirical gap that intuitively does not appear very wide, was executed with some difficulty, particularly in relation to coordinating the independent theoretical stimulus concept with new experimental operations While this extension from conditioning to discrimination learning was taking place, discrimination theory was simultaneously trying to bridge a much wider gap existing between discrimination learning and concept learning Without intending to be critical, this theoretical strategy may be likened to a battle plan in which an attack on an enemy's position is planned from a position not yet won In any case it is necessary to recognize that important defining characteristics of concept learning tasks have emerged from the methods of discrimination learning and the theory of conditioning The main difference between the typical experimental methodology used by S—R psychologists to study discrimination learning and concept learning is that in the former situation single stimulus events are discriminated from each other, while in the latter situation classes of stimuli are discriminated This has led T S Kendler to define the present area of inquiry in the following manner "Concept formation is taken to imply the acquisition or utilization, or both, of a common response to dissimilar stimuli It is the problem of those who study concept formation to analyze the process and determine which variables influence it' (T S Kendler, 1961, p 447)

This definition seems sensible and appropriate It should be recognized, however, that such a definition has its roots in the conditioning generalization extinction model of discrimination learning It is so much a function of these orienting attitudes that it practically forces one to consider concept

learning as continuous with discrimination learning By this definition the simplest kind of concept learning would be exemplified by stimulus generalization, the phenomenon in which a common response is made to *different* stimuli on the same physical dimension

The definition under consideration seems so intuitively obvious that it becomes difficult to think of any alternative If, however, one shifts his frame of reference to that of Piaget's work with concepts (1953), it soon becomes evident that a definition that emphasizes "a common response to dissimilar stimuli" scarcely seems relevant Admittedly the complexity, or perhaps the obscurity, of Piaget's work precludes any satisfactory working definition, but the fact remains that such concepts as *object, space, time, coordination,* and *causality,* concepts which interest Piaget, do not fall comfortably into the aforementioned definition

Now that the orienting attitudes and prejudgments of S—R psychologists who have adopted a conditioning generalization-extinction model have been explicated, it becomes possible to examine their implications for the study of concepts For the present the analysis will be limited to only two points

The first is that major attention has been paid to the learning of concepts, or what has been more commonly referred to as *concept formation, concept attainment,* or *concept identification* This problem has been mainly investigated with an experimental methodology similar in design to that used in discrimination studies, except for convenient modifications borrowed from rote learning procedures (e g , Oseas & Underwood, 1952) As a result such problems as concept utilization, concept modification, level of abstraction of concepts, to name a few, have largely been ignored

The second problem raised is that of the relationship between concept behavior and other forms of responses Operationally specifying a concept as a common response to dissimilar stimuli fails, as already mentioned, to distinguish it from ordinary discriminations and responses to generalized stimuli. Osgood (1953) has tried to make such a distinction, however, not on empirical grounds, but instead in terms of a theoretical principle He raises the question whether Hull (1920), in his classical study in which he found concept attainment resembling ordinary discrimination learning, was " actually studying *concept* formation" His argument is that Hull s *Ss* were required to discriminate a *common* element in a group of Chinese characters These elements were the same in all the characters the group to which the *Ss* had to learn a common response Does not a dog, asks Osgood, go through the same task when he selects out the relevant tone from a multiple-stimulus environment? Osgood prefers to describe Hull's concept learning experiment as a study in labeling, believing that *concepts* require some "abstraction" process His position is expressed in the follow-

ing quotation from his analysis of Fields' (1932) work on the development of the "concept" of triangularity by the white rat

"Patterning his work directly on Hull's approach, he showed that rats could jump toward a triangular form, when paired with other forms in a discrimination situation, despite large variations in size, shading, position, amount of outline, and so forth Yet should we conclude that the rat can understand the *abstract* concept of triangularity? Would the rat respond positively to three dots in a triangular arrangement versus four dots in a square? Or react positively to three people, three places on a map, a three cornered block, as 'triangles'?' (Osgood, 1953, p 667)

The distinctive feature of a concept for Osgood is that it depends on a mediational process That is, instead of a common stimulus in different stimulus patterns becoming associated with a common response, as occurred in Hull's experiment, Osgood believes that a *true* concept emerges when varying stimulus patterns, *not necessarily containing any features in common,* elicit a common mediating response which serves as the cue for conceptual behavior

This apparently meaningful distinction becomes less meaningful under close scrutiny Although it is easy to specify the similarity between two stimulus patterns, it is extremely difficult to decide when two particular patterns fail to possess a feature in common A coffee can and doughnut possess roundness What in this world is completely different from either? A shoe? It is definitely not round but it does possess physical characteristics in common with the coffee can and doughnut They all have weight and substance They all are *something* If one argues this way, then only nothing is uniquely different from something But even here an objection might be raised by some philosopher that nothing is something, or at least the concept of nothing depends on the concept of something

Disregarding the complex semantic overtones of Osgood's distinction, it is necessary to realize that it has served as a bifurcation representing the two directions that investigation of concept attainment has taken

One direction has been the discovery of systematic relationships between stimulus events (words, geometrical designs, etc) and a common response Attention in this sort of work is focused primarily on the stimulus response relationship The second direction is represented by the major interest being directed at the mediational mechanism responsible (at least in the theoretical sense of "responsibility") for concept behavior The internal cue, instead of the association, becomes the main focus of attention An example of the former would be the work of Archer, Bourne, and Brown (1955), Grant, Jones, and Tallantis (1949), and Hovland (1952), who report that certain stimulus features of the experimental task (e g, complexity, distinctive feature, information) are related to the speed of concept

attainment The second orientation is represented by the efforts of Goss (1961) and Kendler and Kendler (1962), who postulate mediating mechanisms and try to determine variables to which they are related

It would be misleading to suggest that the conditioning-generalization-extinction model is the only S—R conceptualization of concept behavior, although as indicated historically it has been the most dominant one In recent years two kinds of phenomena, both of which are represented in S—R language, have been reported that promise to exert influence on the development of research techniques in conceptual behavior One is an outgrowth of operant conditioning techniques, and the other is that of clustering effects in free recall

Operant conditioning —Concept learning tasks that emerge from operant conditioning techniques appear not too different from those which derive from classical conditioning Green (1955) instructed college students to tap a key when a correct discriminative pattern was present, and not to respond when the incorrect pattern was shown The procedure differed mainly from conventional coocept learning methodology involving discrete trials, in that fixed ratio schedules were used in which discriminative stimuli were showo for different periods of time, allowing the *S* to respond repeatedly during one stimulus exposure The results showed that, "The extent to which *S*s discriminated was related inversely to the ratio of responses to reinforcement (1 1, 15 1, and 30 1) and directly to the length of time the discriminative stimuli were presented to them in conditioning (3, 30, and 60 sec)" (Green, 1955, p 180) Of greater interests, and probably of more importance than either the operant conditioning procedure or the results it produced, was the introduction of a verbalization technique which required the *S* to state what he believed to be the characteristics of the correct card This made it possible to study simultaneously the performance of two operant responses, one verbal, the other instrumental key tapping The results showed the correct verbalizations were " related to the parameters of the experiment in the same way as was the key-tapping discrimination" (Green, 1955, p 180)

This dual operant response technique was refined and extended later by Verplanck and Oskamp (1956) in a card sorting task The *S* was required to state a hypothesis before each card placement, with either the hypothesis statement or card placement responses being reinforced One of the most interesting findings was that a discrepancy often existed between the *S*'s stated hypothesis and where he actually placed a card This dissociation between words and action is also noted by Kendler and Kendler (1962) who postulate that several ' horizontal ' chains of S—R associations may be operating simultaneously in concept-learning tasks These chains can function independently of each other, or they can become integrated

by a stimulus from one becoming associated with a response from the other Presumably, when *S*'s verbalizations do not match his actions, the verbal and card-placement chains function in a parallel fashion, i e , they do not interact Combining developmental studies with a mediational S—R model, the Kendlers have sought to understand how the acquisition of verbal processes influences conceptual behavior, i e , how verbalizations interact with instrumental responses

Although Verplanck and the Kendlers share interest in a common problem, their methods of coping with it are decidedly different Functioning within a Skinnerian factual, atheoretical framework tends to encourage treating "verbal hypotheses" and 'card placements" as independent events Each being a response in its own right, the problem is to discover how their rate of responding is influenced by various schedules of reinforcement The problem of the relationship between the two is gingerly touched upon by showing how they covary Such an approach can easily degenerate into a kind of psychophysical parallelism in which the occurrence of each response is scrutinized carefully, while the basic problem of *how* and *why* they interact, or fail to interact, is ignored It must be obvious that the writer's preference is to deal with this problem in a theoretical manner by postulating the mechanisms responsible for the interaction between verbalizations and discrimination responses and then test their implications

Clustering —Although the phenomenon of clustering has emerged from studies in memory, its implications are relevant to conceptual behavior Bousfield (1953) found that when *S* is instructed to recall a list of randomized words in the order that it occurs in memory, he typically reports clusters of related words For example, a list of 40 randomly distributed words comprising 10 names of animals, 10 vegetables, 10 professions, and 10 articles of clothing will not be recalled in a random manner In free recall there will be a tendency for words from the same category to cluster together A name of an animal will have a greater probability of being followed by another name of an animal than by the name of a vegetable, profession, or an article of clothing

Explanations of clustering have varied from those attributing it to the operation of simple associative connections (Jenkins, Mink, & Russell, 1958) to those postulating high level mediated superordinate responses (Bousfield & Cohen, 1956) Whatever the explanation may be, the phenomenon of clustering illustrates concepts in action Words are spontaneously organized into conceptual groups in the absence of any specific instructions Thus clustering would seem to be an important factor in problem solving when the selection of an appropriate word or words has functional significance Judging from the work of Maltzman (1960) on originality as mediated by word associations, work on clustering should ultimately

reveal important information on how concepts are utilized in intellectual behavior

Classificatory behavior has been approached in ways other than those already described In fact, there is practically no limit to the number of pretheoretical models and orienting attitudes that can be used to investigate conceptual behavior In judging what appears to be the most influential approaches, not so much in terms of their achievements, but rather in their hopes for the future, consideration of Piaget, "computer simulation" and mathematical models seems to be in order These will now be discussed in a brief, and hopefully, a cohesive manner, so that the current scene does not appear, as it so often does, as a group of isolated efforts

Piaget

If we dispense with evaluation and analyze Piaget's method of investigating concepts, we find an approach that is markedly different from that used by most S—R psychologists Piaget (1953) does not have any preconceived notion that concept formation is like any other behavior, as was the case for the preconceptions of S—R psychologists of the similarity between classificatory behavior and conditioning and discrimination learning Piaget nevertheless prejudges issues by using a frame of reference to describe behavior that emanates mainly from logic and to a lesser extent from biology As a result he sees conceptual behavior largely in terms of logical operations

Piaget superimposes his logical model on genetic data He attempts to distinguish in a normal child s development successive stages of thought, each of which is characterized by different kinds of logical operations Because of this genetic orientation concepts are considered to emerge at particular ages (just like teeth), with little or no consideration being given to specific learning experiences The major attention then is paid not to the acquisition of concepts but rather to their utilization

Piaget is a psychologist in the European tradition (although there are signs that, like everything else, he is becoming more "American"), and his studies and those of his co-workers, dealing with physical and mathematical problems, do not achieve standards of experimental control and statistical sophistication demanded by that new breed of behavior scientist, the psychonomist Great trust is placed in the verbal responses of children to the questions of the experimenter when it is not always clear that agreement exists about meaning of terms

If substantive issues are put aside, we find that for the most part Piaget's orientation is not contradictory, but is instead supplementary to that of the S—R psychologists The principal area in which this is particularly true concerns their respective treatment of developmental and learning proc-

esses A common criticism, and a justified one, that is often directed at Piaget is that he ignores the influences of learning in the development of conceptual behavior It would be equally appropriate to direct the same kind of criticism, in a reverse form, at the S—R psychologists With comparable indifference they have ignored developmental factors This equal but different neglect stems from a tragic trend throughout the history of psychology to view developmental changes and learning as antithetical processes, a consequence, no doubt, of all the confusion surrounding the heredity-environment controversies What actually is needed to break down this wall dividing these two research areas is some representation of behavior that can be applied to results of developmental as well as learning studies One such attempt that is now being made in a very modest fashion is the analysis in S—R terms of the classificatory and problem-solving behavior of children of different ages (Kendler & Kendler, 1962) Needless to say, it would be helpful to the entire study of psychology if additional attempts were made with different pretheoretical models

Piaget's work also supplements the S—R approach by providing interesting areas of investigation for possible exploitation Although problems of studying such concepts as transitivity, probability, and causality are technically more difficult, there is no reason to believe that they cannot be studied in a manner consistent with behavioristic tradition

Computer Simulation of Cognitive Processes

Perhaps because psychology is reaching its adolescence, it is succumbing more and more to changes in styles of research as well as theorizing Witness the impact of such recent innovations as information theory, mathematical models, and computer simulation of behavior One of several unfortunate consequences of an atomic holocaust would be that we would be deprived of learning which approach was a fad and which had permanent value One can be spared from any doubts about the contributions of computers if one is willing to accept Newell and Simon's (1961a, 1961b) enthusiastic appraisal of simulation of human thinking by computers A behaviorist might find it somewhat difficult to accept these authors' omniscience He will discover that in spite of being awarded their seal of approval for his rigorous methodology, their opinion is that be is interested in insignificant questions

If computer models are to achieve the goals that their enthusiasts have set for them, much greater attention will have to be paid to the behavior that is being simulated and the generality of computer programs The writer is reminded of the reaction of a minister to a lengthy account of the kinds of human behavior that computers have simulated His comment was, "I ll take this stuff seriously when they get a computer that starts

worrying after its parts begin to wear out" This anecdote nicely illustrates the point that in evaluating computer simulation one must also consider the behavior that is being simulated The data that come from the continuous verbal reports of *S* as he is trying to solve a problem may not, as is assumed, mirror the process For example, the reporting might conflict with the thinking and thus change it According to Watt and Ach (Boring, 1929), a person's thinking occurs *before* he knows about what he is thinking If this is true, then at its least a continuous verbal report slows thinking down, and at the very most, actually distorts it Obviously much more information is needed about the relationship between utilizing concepts in problem solving under "normal" conditions and those of continuous verbal report But even if a great difference were found between these two kinds of problem solving, the ability of computers to explain thinking while talking would still be a major contribution Although it would not explain all forms of thinking, it would at least provide insights into the behavior of that small segment of our population who are required to talk while they think (e g, professors who are unprepared for their lectures and wives who never shut up)

More important, perhaps in the long run, is the manner in which these verbal reports are gathered It is awfully easy to shape verbal behavior (e g, Greenspoon, 1955) Unless the most rigorous experimental techniques and controls are used, it is quite possible that experimenters can shape the *S*'s introspective report to fit the mold of some preconceived hypothesis When it is demonstrated that human problem solving resembles that of the computer, it would be wise to ask whether such results are due to the computer simulating human behavior, or whether the experimental technique shaped the subject's behavior so that it resembled that of the computer' (H H Kendler 1961 p 190)

Of course, there is no reason why computer programs have to be applied to verbal reports They can and have been applied to experimental results such as those from partial reinforcement and rote learning studies (Newell & Simon 1961b) The important point is that to apply a program to a single experimental result is only the beginning of effective theorizing What are needed are general theories—or programs—that can be applied successfully to a variety of empirical data To return to Newell and Simon's General Problem Solver a computer program that fits some logical manipulations (the fit is not perfect since the human leaves out some of the steps of the computer) the giant step forward will come when it, or some other program is expanded to fit a variety of demonstrated phenomena such as functional fixedness the relative effectiveness of reversal and nonreversal shifts in conceptualization for different age groups, the effects of verbal pretraining on subsequent concept behavior, etc

Lest these comments be interpreted as completely negative, reference is made to Hovland (1960), who appropriately points out that computers can function as potential *aids* " in sharpening our formulations concerning mental processes and phenomena," in encouraging " theories that have both descriptive and predictive power," and for coping with the complex problem of dealing with a multitude of interacting variables

Mathematical Models

Formally speaking, there is no difference between representing behavior by computer programs or by mathematical models A computer program " used as a theory has the same epistemological status as a set of differential equations or difference equations used as a theory' (Newell & Simon, 1961b, p 2013) But in other respects they are different Although it is difficult to generalize from all the various mathematical models, it is safe to say that there is at least one difference that is relevant to this discussion To an appreciable extent mathematical models, unlike much of the computer simulation work, have emerged from the S—R language and methodological traditions

One effort to deal with concepts has taken the form of a concept identification model (Bourne & Restle, 1959) which is an elementary extension of a mathematical model for discrimination learning (Restle, 1955) The reason that this transition is made with relative ease, aside from the clarity of the model, is that Bourne and Restle restrict their attention to relatively simple conceptual behavior having its roots in empirical relationships between clearly defined stimulus variables (e g, number of relevant dimensions, number of irrelevant dimensions, delayed feedback) and choice behavior Their work can be located conceptually along the 'nonmediational" path emanating from the bifurcation that has been associated with Osgood's distinction between labeling (nonmediational) and concept behavior (mediational)

Mathematical models lend themselves to this sort of single stage S—R analysis Problems that have required a mediational analysis tend to be avoided When they have been dealt with, as is the case of Restle's analysis of learning set data (1958), mediating stimuli resulting from responses of the organism are postulated It would be most interesting if attempts are made to deal with 'mediational type" phenomena with a single stage S—R mathematical model For example, can Restle's discrimination theory (1955) with the assistance of his 'adaptation' mechanism represent the changes in the relative rapidity in which reversal and nonreversal shifts are executed by children of different ages (Kendler & Kendler, 1959, Kendler, Kendler, & Wells, 1960)?

There is little doubt that mathematical models will play a more and

more important role in attempts to investigate and interpret classificatory behavior as research moves ahead Disagreemens exist as to what exactly this role will be Will it be an innovator that leads to the discovery of new phenomena while systematizing the old, or will it make its main contribution only in terms of the latter function? In either case, the warning of Skinner that comes from the wings, or from above, depending how one views the current sceoe in psychology, should not be *completely* ignored, "Wbat is needed is not a mathematical model, constructed with little regard for the fundamental dimensions of behavior, but a mathematical treatment of experimental data" (Skinner, 1961, p 62)

The Nature of Concepts

The impression has no doubt been given of talking around the concept of the concept Any reader, hoping to discover what a concept really is, would no doubt bave his curiosity unsatisfied, if not downright frustrated Admittedly, the writer or psychology may be partly at fault, but it can be argued, perhaps in self defense, that the major difficulty lies with this hypothetical reader's uncritical curiosity After all, a concept is not a single thing or event whose attributes and functions can be simply listed and described A concept is a complicated psychological pbenomenon, a complete description of which will be contained someday in a theory capable of explaining the numerous empirical laws involving the term *concept* That day will only come about when knowledge about *all forms* of behavior is greatly increased

Although the question of the discontented reader cannot now be satisfactorily answered, the direction whicb such ao answer will take can be suggested In order to do this, some method of conceptualizing behavior in general, and classificatory behavior io particular, must be offered Stimulus response language can serve this function As a *first approximation*, concepts bave three conceptual properties they are associations, they function as cues (stimuli), aod they are responses This classification scheme, whose divisions are not as ironbound as they should be, is useful in ordering some of the major empirical problems existiog in cooceptual behavior

Concepts as Associations

The greatest experimental effort has been directed toward the learning of concepts Typically this has meant the association between dissimilar stimuli and a common response The experimental methodology used to investigate this process resembles in many ways the conventional discrimination procedures used with animals, a single stimulus event is presented,

whether it be in a card sorting task or a rote learning type experiment, and the *S* is required to make a response which in turn is either reinforced or not One important problem that has been raised is whether the concept learning that goes on in experiments of this sort really represents a single kind of concept learning or instead certain stages of concept learning The terms *concept identification* and *concept acquisition* (or attainment) highlight a possible distinction In a conventional card sorting task the *S*, who is usually a college student, does not 'acquire a concept such as number or shape or color Instead he identifies one Prior to training the *S* knows these concepts, he merely has to learn appropriate responses to the appropriate stimuli But the basic associations have been formed His task is different from that of a child who does not know what shape or color or number is at the beginning of the experiment The child has to learn, for example, the *associations* between square shapes and a common response and circles and another response This distinction is operationally meaningful if we define concept identification tasks as occurring when instructions could produce the same behavior (e g, sorting in terms of number) as the conventional training procedures Concept acquisition would be restricted to the situation in which a simple set of instructions would not suffice The *S* would have to acquire a concept from the very beginning"

This distinction does not rest only upon the ages of the *S*s or their verbal ability College students who could learn to *identify* a number concept would have to learn to *acquire* such concepts as *disinhibition* and *spontaneous recovery* And correct verbalizations of children in simple concept learning tasks are not always a guarantee that their choice behavior will be appropriate (Kendler & Kendler, 1962)

It would seem that acquisition occurs prior to identification, although there is no reason to suggest that both processes could not occur simultaneously The basic problem is to discover how, if at all, these processes differ Is only the acquisition process truly associative, while the identification process is a relatively simple case of transfer, in which a new response is associated to an old class of stimuli? Or are these processes extreme points on a single dimension? If comparable studies could be done with concept identification and concept acquisition tasks, then the differences between the two processes would be better understood

The selection of a particular experimental procedure would be a problem for the investigator who decided to try to answer the questions just raised Concept learning experiments have barely tapped the wide variety of possible conditions—as well as questions to ask Answers to the simple practical question of what are the optimal conditions for concept learning can reap a rich harvest This problem has not been ignored, as indicated by interest shown in the relative importance of positive and negative instances

(Hovland, 1952), in the relative difficulty of identifying conjunctive and disjunctive concepts (Bruner, Goodnow, & Austin, 1956), and the retarding effects of irrelevant information (Archer, Bourne, & Brown, 1955) But obviously much more has to be done A particularly important problem is suggested by a study of Marx, Murphy, and Brownstein, who in seeking to train Ss to recognize classes of complex stimuli find that training with ' certain kinds of abstracted patterns results in greater recognition than presentation of the fully drawn stimuli" (1961, p 459) Although their study is primarily concerned with what may be properly described as principles of perceptual organization, the problem touches upon one that is central to conceptual behavior, whether it be concerned with classifying geometrical patterns, words, ideas, or what have you This problem is what kind of learning allows for the most effective transfer from the abstract to the specific and from the specific to the abstract

In discussing associative problems of conceptual behavior the topic of retention cannot be ignored We know concept identification occurs more rapidly if the *S* is freed from the task of remembering previous correct and incorrect instances (Bruner, Goodnow, & Austin, 1956) But if learning a concept is dependent to some extent on retention, as it usually is, what are those principles of retention that can be exploited to facilitate concept learning? It is customary when studying concept learning with memory drums or card sorting tests to present all the positive and negative instances in successive mixed sequences Although these sequences are necessary for experimental control, the order itself may be an important variable Perhaps an instance should be repeated until the correct response is forthcoming The problem being raised is related to that old and seemingly forgotten one of ' part ' versus ' whole ' learning Previous attacks on this topic failed to provide any definitive answer Perhaps psychologists did not know enough about learning and methodology at that time But it might be noted that problems of conceptual behavior were largely ignored when this topic was originally investigated An approach that considers conceptual behavior might make sense out of what has appeared for many years to be nonsense

It is possible to make the transition between associative and stimulus and response problems of conceptual behavior by raising questions related to developmental factors in concept learning What is the psychological basis of the developmental changes occurring in various kinds of concept learning? To attribute these differences to intelligence, at best, is offering an incomplete answer and, at worst, is begging the question Is the locus of these differences in associative functioning or stimulus or response mechanisms? This question cannot be answered now (a number of partial answers are being offered), but if stimulus response language is useful, some answer

should be forthcoming after appropriate research is completed And these answers will represent giant strides forward in understanding classificatory behavior

Concepts as Cues

One of the major differences between S—R correlationists and mediational theorists is that the former are interested in concepts as associations and the latter in concepts as cues The result is that the former have been more concerned with concept formation, while the latter are primarily interested in concept utilization

Although much theoretical use has been made of concepts as cues (Goss, 1961, Kendler & Kendler, 1962), there are some (Bruner, Goodnow, & Austin, 1956, Miller, Galanter, & Pribram, 1960, Shepard, Hovland, & Jenkins, 1961) who believe the mediational approach with its emphasis on the cue function of concepts is not enough They argue that something like "ideas" or "principles" or "strategies" are needed in addition to what is referred to as S—R associations This difference reminds one of what appeared to some to be the basic issue surrounding the latent learning controversy, namely, whether S—R associations or cognitions were learned This distinction has always puzzled the writer (Kendler, 1952) because it seems to reflect personal preference for models and language systems adopted to represent behavior instead of fundamental theoretical assumptions If the vast potentialities of S—R language resulting from chaining and mediated generalization of verbal behavior are considered, then what some prefer to call "principles" or "ideas" or "strategies" can be, although need not be, expressed in terms of cues arising from *chains* of S—R associations

A somewhat more meaningful aspect of this problem was expressed in an exchange between Bousfield (1961) and Osgood (1961) concerning the justification of postulating a concept of "meaning" that is separate and independent of the cue functions of words Bousfield is willing to dispense with "meaning" because it is unnecessary and it generates confusion Meaning for him can be handled by word associations, a concept's meaning would be expressed by the word it elicits Osgood disagrees, believing that responses to words depend on other things besides meaning (e g , syntax) Meaning per se is better measured by the semantic differential The reaction to this disagreement by the members of the symposium at which this exchange occurred suggested that this issue was not crystal-clear It may be that when we expand our method of measuring the cue function of words to include chains of words (i e , phrases and sentences), then this apparent difference will disappear Or it may be, as Osgood argues, that meaning can never be reduced to the cue function of concepts Whatever the outcome of this intratheory squabble, the results cannot help

but bave important repercussions fnr mediational theory in general and conceptual bebavior in particular

The mechanism basic to both these intertheoretical and intratheoretical differences can be categorized as problems of concept generalization Having learned a conceptual response, what are the instances that will evoke it? A question like this is extremely relevant to the area of problem solving Perhaps when programed learning is used more as an experimental tool (Gagne & Brown, 1961) than as a sales commodity, basic information about concept generalization will be obtained

Concepts as Responses

When a child bas acquired a concept, what response has he learned? According to current mediational theories (Goss, 1961) he has acquired some implicit response, usually, although not necessarily, verbal in nature, which serves as a cue for his overt behavior This is obviously not the only possible interpretation Zeaman takes a position more in line with Gibson's (1959) perceptual formulation, " distinctiveness is acquired not by the addition of response produced interoceptive stimulation but rather by the acquisition of responses (observing or attention) whose function is to make effective certain features of *exteroceptive* stimulation not previously responded to ' (Zedman, personal communication)

This alternative to S—R mediational theory should—and no doubt will—be pursued But certain problems must be recognized First, it will not be sufficient to equate observing responses with attention Although they are functionally equivalent in that they both operate to "select out" from the total pattern of stimulation those components that will become associated, it may be that principles governing their operation are different Observing, or what some prefer to call receptor-orienting responses, will determine what part of the environment will strike the organism s sensorium Attention, on the other hand, decides what stimulus component of a pattern of stimuli, falling within the "receptor gaze" will stand out and become associated In short, both observing and attending will influence the stimuli that are to be associated, but their influence may operate through different mechanisms, the former through principles governing the learning and performance of instrumental responses, the latter through principles governing perceptual organization

The second problem that must be recognized in developing the perceptual formulation of the sort Zeaman suggests is to understand how it differs from a mediational approach There seem to be three possibilities One is that perceptual and mediational responses can be parts of separate segments of the same behavioral chain (Kendler, Glucksberg, & Keston, 1961) Another possibility is that perceptual and mediational responses

interact, thus influencing each other in a reciprocal manner Finally, they may not be distinct events even though they have been assigned different names It would surprise the writer if all three possibilities did not possess some degree of validity for at least some situations For the moment it would be wise to consider the possibility that apparent differences could hide intrinsic similarities

Disregarding the relationship between mediation and perception, it would seem reasonable to expect that many, although not necessarily all, developmental differences will be discovered to be localized in different response capabilities This would be true for many of the differences noted in athletic skills, and there is probably some psychological similarity between athletic and intellectual skills Response differences resulting from developmental changes in human intellectual functioning can be categorized in four different groups First, the degree of abstraction can be increased Second, partly as a result of language development, the child can learn to respond in a mediated manner Third, the acquired verbal habits increase in complexity and uniqueness One reason that Gagne and Brown (1961) find that self discovery of a principle is superior to being taught the principle is that the former procedure allows the principle to fit better into the individual's existing verbal system Fourth, as development proceeds, the verbal behavior which initially develops as an independent chain has increased opportunity to become integrated with other patterns of behavior that are constantly being learned

Before terminating this section attention should be called to the importance of conceptual behavior in education If researchers in the field of programed learning believe that the major response being learned is the one the student writes in the frame, they are sadly mistaken (Gagne & Brown, 1961, Kendler, 1959b, Silverman, 1961) What is being learned are concepts, and what is modified are large segments of the student's verbal repertoire Unless we understand this process better than we do now, the hopes of programed learning will never be realized

THE PLACE OF CLASSIFICATION LEARNING IN THE PSYCHOLOGY OF LEARNING

In a symposium such as this one cannot avoid considering the interrelationships among the various categories of human learning In fact, the present sequence of topics implies an organization of learning processes that is widely accepted The underlying theme of this organization is that conditioning is the simplest form of learning One can observe in the conditioning situation the formation and strengthening of associations in all their natural glory and simplicity As one proceeds from conditioning to

rote learning through probability learning and finally on to problem solving, the learning phenomena become more and more complex, more and more a function of a larger number of variables and the interrelationships among them This organization of the psychology of learning from the apparently simple to the complex is so eminently reasonable and sensible that it is often accepted as valid In fact, the writer (Kendler, 1959a) once directed a criticism toward statistical learning theorists for using such complex empirical phenomena as probability learning to arrive at their fundamental theoretical assumptions It was suggested that the appeal of probability learning was not its psychological simplicity, but instead the quantitative simplicity of the data it produced This criticism is justified if one accepts the idea that conditioning mirrored most clearly the fundamental processes of learning Such a "conditioning-oriented" view is obviously not demanded by the facts, but instead represents a prejudgment as to how the facts will ultimately be ordered

Realizing this point, it is as justified to organize all learning around the facts of probability learning as around the facts of conditioning In short, the answer to questions concerning the interrelationships among various categories of learning must of necessity be a theoretical answer This does not mean that one answer is as good as another Instead it means that the best answer must be supported by the best theory

Although it might be possible to argue that concept learning should be considered as the core of all learning, the writer would just as soon not defend this view But it might be mentioned that an allied view proposed by Tolman (1932) that discrimination learning is basic to all learning be considered more seriously It is impossible to consider any learning situation in which competition between responses is completely absent In classical conditioning the competition between the conditioned response and other reactions (e g, the "investigatory" response) are typically ignored The conditioned response, as a result of a long series of acquisition trials, may become so powerful that evidence of competing responses is absent. But the introduction of experimental extinction procedures will usually show that hidden behind every dominant response is another response that can take over

When viewing the history of the psychology of learning over the past few decades one can notice a shift of interest from the development of a single association to the competition between associations Perhaps this shift represents a reorientation toward a more fundamental problem—or at least to a problem that is equally fundamental The implication of this is that if the various categories of learning are organized into some meaningful relationship, perhaps discrimination learning with its emphasis on the process of habit competition, should be assigned the central role

Concluding Comments

An attempt has been made to show how investigations of conceptual behavior have been shaped by preconceptions of such behavior Some of the more popular preconceptions have been reviewed and analyzed These models tend to focus attention and obtain information in different problem areas The analysis was terminated by attempting to systematize the problems of conceptual behavior in terms of three psychological functions of concepts resulting from a stimulus response analysis, and pointing out that the location of conceptual behavior against a background of all psychology will ultimately depend on the theory of behavior which is accepted

It is easy to get depressed about the current status of our knowledge of conceptual behavior We know so little We have so much to learn Yet depression, or even disappointment, is not justified Different empirical techniques have been developed, knowledge is being gathered, and some integrating ideas, even if they are limited in scope, are being offered No doubt individual psychologists viewing the effort that is being expended by behavior scientists in this field could recommend a more economical and strategic approach For example, the standardization of our research techniques could be improved This will be achieved not by wishing or proclamation, but by the development of techniques that capture the imagination of the younger psychologists who enter this fascinating field Criticism about our present concepts of conceptual behavior can be offered They are both clumsy and confusing But problems of definition will dissolve as the field matures The great need is to get more researchers in the field of conceptual behavior And this is exactly what seems to be happening

References

Archer, E J, Bourne, L E, & Brown F G Concept identification as a function of irrelevant information and instructions *J exp Psychol*, 1955, *49*, 153–164

Bergmann, G Psychoanalysis and experimental psychology A review from the standpoint of scientific empiricism *Mind* 1943, 52, 122–140

Boring, E G *A history of experimental psychology* New York Appleton-Century, 1929

Bourne L E & Restle, F Mathematical theory of concept identification *Psychol Rev*, 1959, 66, 278–298

Bousfield W A The occurrence of clustering in the recall of randomly arranged associates *J gen Psychol* 1953 49, 229–240

Bousfield, W A The problem of meaning in verbal learning In C. N Cofer (Ed.), *Verbal learning and verbal behavior* New York McGraw Hill, 1961 Pp 81–91

Bousfield, W A & Cohen, B H Clustering in recall as a function of the number of word-categories in stimulus word lists *J gen Psychol*, 1956 54, 95–106.

Bruner, J S, Goodnow, J J, & Austin, G A. *A study of thinking* New York John Wiley, 1956

Dodwell, P C Coding and learning in shape discrimination *Psychol Rev* 1961, 68, 373–382

Ehrenfreund, D An experimental test of the continuity theory of discrimination learning with pattern vision. *J comp physiol Psychol*, 1948, 41, 408–422

Estes, W K. The statistical approach to learning theory In S Koch (Ed.) *Psychology A study of a science* New York McGraw Hill, 1959 Vol 2, pp 380–491

Fields P E. Studies in concept formation I The development of the concept of triangularity by the white rat. *Comp Psychol Monogr*, 1932, 9, Whole No 42

Gagné, R M, & Brown, L. T Some factors in the programing of conceptual learning *J exp Psychol* 1961, 62, 313–321

Gibson, J J Perception as a function of stimulation In S Koch (Ed) *Psychology A study of a science* New York McGraw Hill, 1959 Vol 1, pp 456–501

Goss, A E Early behaviorism and verbal mediating responses *Amer Psychol*, 1961, 16, 285–298

Goss, A. E Verbal mediating responses and concept formation *Psychol Rev*, 1961, 68, 248–274

Grant, D A., Jones, O R., & Tallantis, B The relative difficulty of number, form, and color concepts of a Weigl type problem. *J exp Psychol*, 1949, 39, 552–557

Green, E. J Concept formation A problem in human operant conditioning *J exp Psychol* 1955, 49, 175–180

Greenspoon, J The reinforcing effect of two spoken sounds on the frequency of two responses. *Amer J Psychol*, 1955, 68, 409–416

Hanson, N R. *Patterns of discovery an inquiry into the conceptual foundations of science* New York Cambridge University Press, 1958

Hernandez Peon, R Scherrer, H, & Jouvet, M Modification of electrical activity in cochlear nucleus during "attention" in unanesthetized cats. *Science*, 1956, 123, 331–332.

Hilgard E. R., & Marquis, D G *Conditioning and learning* New York Appleton Century, 1940

Hovland, C. I The generalization of conditioned responses. I. The sensory generalization of conditioned responses with varying frequencies of tone *J gen Psychol.*, 1937, 17, 125–148

Hovland C. I A "communication analysis" of concept learning. *Psychol Rev* 1952, 59, 461–472.

Hovland, C. I Computer simulation of thinking. *Amer Psychologist*, 1960, 15, 687–693

Hull, C. L. Quantitative aspects of the evolution of concepts *Psychol Monogr*, 1920 28, Whole No 123

Hull, C. L. Knowledge and purpose as habit mechanisms. *Psychol Rev* 1930, 37, 511–525

Hull, C. L. *Principles of behavior* New York Appleton-Century, 1943

Jenkins J J Mink, W D., & Russell, W A. Associative clustering as a function of verbal associative strength. *Psychol Rep*, 1958, 4, 127–136

Kendler, H H "What is learned?"—A theoretical blind alley *Psychol. Rev*, 1952, 59, 269–277

Kendler, H H Learning. *Annu Rev Psychol.* 1959, 10, 43–88 (a)

Kendler, H. H. Teaching machines and psychological theory In E. Galanter (Ed.)

Automatic teaching: The state of the art New York Wiley, 1959 Pp. 177–186. (b)

KENDLER, H. H. Problems in problem-solving research. In *Current trends in psychological theory.* Pittsburgh. University of Pittsburgh Press, 1961. Pp 180–207.

KENDLER, H. H., & KENDLER, T. S. Effect of verbalization on reversal shifts in children. *Science,* 1961, 134, 1619–1620.

KENDLER, H. H, & KENDLER, T. S. Vertical and horizontal processes in problem solving. *Psychol. Rev,* 1962, 69, 1–16.

KENDLER, H. H, GLUCKSBERG, S, & KESTON, R. Perception and mediation in concept learning. *J. exp Psychol,* 1961, 61, 186–191.

KENDLER, T. S Concept formation In *Annu. Rev. Psychol,* 1961, 12, 447–472

KENDLER, T. S, & KENDLER, H. H. Reversal and nonreversal shifts in kindergarten children. *J. exp Psychol,* 1959, 58, 56–60.

KENDLER, T. S, KENDLER, H. H, & WELLS, D Reversal and nonreversal shifts in nursery school children. *J. comp. physiol. Psychol,* 1960, 53, 83–88

KOCH, S. Clark L Hull. In W. K Estes, et al *Modern learning theory.* New York: Appleton Century, 1954. Pp 1–176.

KOFFKA, K. *Principles of Gestalt psychology.* New York. Harcourt Brace, 1935.

KRECHEVSKY, I. An experimental investigation of the principle of proximity in the visual perception of the rat. *J. exp. Psychol,* 1938, 22, 497–523.

LACHMAN, R. The model in theory construction *Psychol. Rev,* 1960, 67, 113–130

LASHLEY, K. S The mechanism of vision XV. Preliminary studies of the rat's capacity for detail vision. *J. gen. Psychol,* 1938, 18, 123–193.

MALTZMAN, I. On the training of originality. *Psychol Rev,* 1960, 67, 229–242

MARX, M. H, MURPHY, W. W, & BROWNSTEIN, A. J. Recognition of complex visual stimuli as a function of training with abstracted patterns *J. exp. Psychol,* 1961, 62, 456–460.

MILLER, G. A., GALANTER, E, & PRIBRAM, K. H. *Plans and the structure of behavior.* New York: Holt, 1960.

NEWELL, A, & SIMON, H. A. In *Current trends in psychological theory.* Pittsburgh: University of Pittsburgh Press, 1961. Pp. 153–179. (a)

NEWELL, A, & SIMON, H. A. Computer simulation of human thinking. *Science,* 1961, 134, 2011–2017. (b)

OSEAS, L., & UNDERWOOD, B. J. Studies of distributed practice. V. Learning and retention of concepts. *J. exp Psychol,* 1952, 43, 143–148.

OSGOOD, C. E. *Method and theory in experimental psychology.* New York: Oxford University Press, 1953.

OSGOOD, C. E. Comments on Professor Bousfield's paper. In C. N Cofer (Ed.), *Verbal learning and verbal behavior.* New York. McGraw-Hill, 1961. Pp. 91–106.

PAVLOV, I. P. *Conditioned reflexes* (Translation by G. V. Anrep). London: Oxford University Press, 1927.

PIAGET, J. *Logic and psychology.* Manchester, Eng: Manchester University Press, 1953

RESTLE, F. A theory of discrimination learning. *Psychol. Rev,* 1955, 62, 11–19.

RESTLE, F. Toward a quantitative description of learning set data. *Psychol Rev,* 1958, 65, 77–91.

ROBINSON, E S. *Association theory to-day* New York: Appleton-Century, 1932.

SHEPARD, R N, HOVLAND, C. I, & JENKINS, H. M Learning and memorization of classifications. *Psychol. Monogr,* 1961, 75, Whole No. 517.

SILVERMAN, R E Response mode, pacing, and motivational effects in teaching machines Technical Report NAVTRADEVCEN 507 3, U S Naval Training Device Center, Port Washington New York, 1961

SKINNER, B F The flight from the laboratory In *Current trends in psychological theory* Pittsburgh University of Pittsburgh Press, 1961 Pp 50–69

SPENCE, K W The nature of discrimination learning in animals *Psychol Rev*, 1936, 43, 427–449

SPENCE, K W *Behavior heory and conditioning* New Haven Yale University Press, 1956

TOLMAN, E C *Purposive behavior in animals and men* New York Appleton Century, 1932

VERPLANCK, W S, & OSKAMP, S Equivalent effects of reinforcing card placement and hypothesis statement in maintaining a concept under partial reinforcement *Amer Psychologist* 1956 **11**, 421 (abstract)

On Verbalizations and Concepts

COMMENTS ON PROFESSOR KENDLER'S PAPER

E. JAMES ARCHER
University of Wisconsin

It is difficult to discuss a paper which is so interesting, agreeable, and informative, without being picayune and trivial.

Professor Kendler has presented the history of the development of the methodological biases in the study of concepts very elegantly. Perhaps his concern with "models of behavior [being] extended *to* concept learning" (p. 212) would be ameliorated if viewed differently. It is not uncommon to hear our nonpsychologist colleagues express amazement over our exhaustive exploration of seemingly trivial and insignificant problems—because for those problems we have a rigorous methodology—and our ignoring of the bigger, obviously important problems—for which we haven't evolved satisfactory techniques as yet. Viewed differently (and maybe a bit smugly), could the S—R association models be seeking justification and merit *because* they seem to be extendable to the high order, complex-real life phenomenon of concept formation?

In any case, no quarrel can be found with Professor Kendler's description of the similarities of the methodologies of conditioning and discrimination learning, and then discrimination learning and concept learning.

I am not sure that I share his concern for the yawning gaps he sees among these three classes of behavior, at least as far as infrahuman animals and nonverbal children are concerned. I share Professor Grant's (p. 28) views on the need for an anthrocentric psychology. The gap between those who verbalize and those who do not seems much more significant.

The three classes of studies: conditioning, discrimination learning, and concept formation seem to lie along a dimension of stimulus and response complexity, but *verbalization* appears to provide a qualitative difference. The recognition of the importance of verbalization is nicely shown by the increased interest in such problems as varied as the effects of attitudes and self-instruction in eyelid conditioning, the effect of verbal pretraining on perceptual motor skills, and strategy selection in concept identification. As Miller, Galanter, and Pribram said, "It is so reasonable to insert between the stimulus and the response a little wisdom. And there is no

particular need to apologize for putting it there, because it was already there before psychology arrived" (1960, p 2) There is even a growing interest in finding out what the subject thinks he is *supposed* to be doing in an experiment

As Professor Kendler pointed out, the current trend in many S—R theories (e g, Spence 1956) is for large parts of the environment to which the *S* is responding to be introjected into the organism If a hungry rat has fractional anticipatory goal responses, I will wager a motivated college sophomore will say to himself, "I'll try the green circles, too, and see if they are right"

I would like to offer some encouragement to Professor Kendler to take issue with the definition of concept formation as "the acquisition or utilization of a common response to dissimilar stimuli" (T S Kendler, 1961, p 447) At least, I would like to promote the position he almost takes when he states that "concepts have three properties they are associations, they function as cues (stimuli), and they are responses" (p 226) There is a fourth property—which subsumes these three—they are *words* More specifically they are meaningful words which label classes of otherwise dissimilar stimuli These words are associations, they are cues, and they are responses The most obvious argument against going this far is to point out that monkeys, for example, can acquire what appears to be a *concept of oddity* and yet, as far as we know, they do not discuss these "words" with one another On the other hand, if the precursor to the *word* is a feeling, an intuition, a hunch, (and all of this is pretty private), then this could account for the exasperating slowness and difficulty of acquisition of such concepts as oddity, double-alternation, transposition and the like in infrahuman and pre-verbal human *Ss* Here I am suggesting that dissimilar stimuli are beginning to be perceived as "going together" via a process of conditioning

This is not proposed as an all-or none phenomenon Concepts are not just words—they are meaningful words used as class labels—and the acquisition of the meaning is surely a continuous and variable process

To use Professor Kendler's example of the distinction between concept identification and concept acquisition, 'College students who could learn to *identify* a number concept would have to learn to *acquire* such concepts as *disinhibition* and *spontaneous recovery*" (p 227) Even if these college students haven t had a course in psychology, I bet they could give some definition of these two terms, but to give the definition in the context we mean would require the acquisition of further and specialized meaning At such time they would, we would say, have acquired the concepts of disinhibition" and spontaneous recovery'

Furthermore I m not sure that the two approaches to concept forma-

tion Professor Kendler described are as different as he suggests On the one hand he sees attention focused on the S—R relationship, e g , the Archer, Bourne, and Brown study (1955), and on the other hand he sees attention focused on mediational processes (Kendler & Kendler, 1962) I wonder how different these really are By varying the amount of irrelevant information we *did* vary the complexity of the stimuli, but just as surely we also varied the number of hypotheses which *S* could consider The amount of irrelevant information was more easily specified than the amount of self-verbalization which was provoked in *S*

While the Grant, Jones, and Tallantis paper (1949) stressed perseverative errors, a fairly objective external kind of response, these errors were surely grounded in *S*'s verbalizations of his strategies While Bruner, Goodnow and Austin (1956) stressed the relative difficulty of conjunctive and disjunctive concepts, these differential difficulties were surely grounded in *S* s ability or inability to verbalize the concepts

Emphasizing the importance of S s verbalization is hardly new (Goss, 1961), but we might find it advantageous to consider the extreme position of insisting that a similarity or relationship be verbalized, at least covertly, *before* it be called a concept And when it is verbalized, we are faced with a rather new and distinct class of variables As is well known (Katona, 1940), something special seems to happen when the *S* "understands' the concept Recalling the example in Underwood s (p 63) discussion of "conceptual similarity and free learning,' when two of four lists of words had four categories of words, there was remarkably better retention of the lists Here obviously, 'something special" happened This 'something special" need not always be beneficial If you recall Professor Kimble s (p 37) reference to Stuart Chase's phrase of the 'tyranny of words, the particular verbalization could affect the structure of the concept, even distort it In our world of ideological conflict, this effect of the verbalization has very significant implications Could not this 'understanding' be coincident with *S*'s ability to correctly verbalize the meaningful label or word which makes the concept a concept?

Problem solving is less adequate when the rules are memorized by rote The rules even show themselves to be subject to interference and forgetting The rules get interchanged or garbled or forgotten completely when the concept has not been attained On the other hand, how much more resistant to interference and forgetting are thoroughly understood concepts! In fact, the concept may remain even after the rote learned rules have been forgotten

Maybe we are too eager when we try to extend principles of conditioning and discrimination learning to concept formation Perhaps there is a set of new independent variables which are being ignored because of

our self-imposed restrictions Perhaps some of these variables have already been suggested by some Gestalt notions of organization

As for Professor Kendler's hopes for Piaget's method, "computer simulation," and various mathematical models, I am less optimistic

Piaget's methods place an extraordinary reliance on the overt verbalizations of his young *S*s In reading over many of the transcripts, one gets an impression that many of the responses are forced The child seems to say things because he is pressed to say something—anything And consequently, some of these highly regarded data records are gibberish Piaget's methods call for some highly skillful filtering and selecting to achieve order

The computer simulation and mathematical model approach may give us some magical moments of understanding, but I suspect that the programs and models will have to sense when the "big step" from pre-verbalization to verbalization takes place and apply the new variables appropriately I am suggesting that there may be a gradual development of attending to stimuli, selection of information, formulation of hypotheses, testing of these hypotheses, identification of relevant and irrelevant information, elimination of redundant relevant information and a gradual but final "firming up" of a verbal statement of the concept Perhaps some of these steps are S—R associations with little mediation but I think it is more probable that each step has associated internalized verbalizations A subject does not operate in a verbal void He keeps talking to himself Most of what he says is irrelevant and/or incorrect But he keeps trying to find the "rule" which will get him out of the experimental "cage" When he finds it, he behaves very differently He doesn't easily forget the rule, in fact he sometimes seems obsessed with its application

It is when this "rule" is found that we seem to need a new class of variables When these are known and understood, then computer simulation will be useful As Professor Kendler states, before it can work, the computer has to know what it is supposed to simulate

In summary, I agree with Professor Kendler about the growing recognition of the role that verbalization plays in concept formation, identification and acquisition I seem only to disagree with him on the pervasiveness and fundamental importance of verbalization to the very existence of a concept

References

Archer, E. J Bourne, L E & Brown F G Concept identification as a function of irrelevant information and instructions *J exp Psychol*, 1955 49, 153–164

Bruner, J S., Goodnow J J & Austin G A *A study of thinking* New York Wiley 1956

Goss, A E Verbal mediating responses and concept formation *Psychol Rev*, 1961, 68, 248–274

Grant, D A, Jones, O R, & Tallantis, B The relative difficulty of number, form, and color concepts of a Weigl type problem *J exp Psychol*, 1949 39, 552–557

Katona, G *Organizing and memorizing* New York Columbia University Press, 1940

Kendler, H H, & Kendler, T S Vertical and horizontal processes in problem solving *Psychol Rev*, 1962, 69, 1–16

Kendler, T S Concept formation In *Annu Rev Psychol*, 1961, 12, 447–472

Miller, G A, Galanter, E, & Pribram, K H *Plans and the structure of behavior* New York Holt, 1960

Piaget, J *Logic and psychology* Manchester, Eng Manchester University Press 1953

Perceptual-Motor Skill Learning[1]

PAUL M FITTS
The University of Michigan

Living, moving, and behaving are almost synonymous terms Thus the study of motor and perceptual motor skill learning is in a very real sense the study of a large segment of the field of psychology However, there are two reasons why skill learning is not an especially appropriate or useful subclass of learning situation First, the theoretical framework within which skilled performance is now being viewed by most students of this topic is such that sharp distinctions between verbal and motor processes, or between cognitive and motor processes serve no useful purpose Second, since the processes which underlie skilled perceptual motor performance are very similar to those which underlie language behavior as well as those which are involved in problem solving, and concept formation, we should expect to find that the laws of learning are also similar, and that no advantage would result from treating motor and verbal learning as separate topics I realize, of course, that the distinction between verbal and motor processes is often a convenient one for practical purposes, and that the distinction has been common in theories of learning For this reason the first part of the present chapter is devoted to a general discussion of present theories regarding the nature of skilled performance and to an effort to establish the close relationship between verbal and motor processes

Interest in skill learning has fluctuated widely over the years, being rather high around the turn of the century and again at mid century, but remaining low from about 1910 to 1940 The most recent resurgence of interest in the topic of skill has resulted chiefly from influences outside of psychology I refer to the development of complex mechanisms for use in control, communication, and computing operations and to the parallel development of theory and mathematical models relating to such processes These models have proven to be very fruitful in stimulating psychological research, and in providing an integrative framework within which the similarities between different aspects of behavior, and various learning

[1] This chapter is based in part on research supported by the U S Air Force, Office of Scientific Research, under Contract No AF 49 (638)-449

situations become more apparent The heuristic importance of these models is further justification for devoting the first part of this paper to a consideration of their implications for learning theory

Definitions and Taxonomy

Definitions

I realize that my view of the area encompassed by the term perceptual motor skill may differ from the views of many students of human learning So a few pages devoted to specification of some defining operations and to the development of background information are warranted

First of all, I intend to emphasize the term *skill*, rather than the terms perceptual or motor By a skilled response I shall mean one in which receptor-effector feedback processes are highly organized, both spatially and temporally The central problem for the study of skill learning is how such organization or patterning comes about

The matter of what muscle groups are involved in a particular behavioral sequence often is quite incidental and certainly no more important than the question of which retinal elements or which segments of the basilar membrane are involved in detecting a stimulus pattern The more relevant issue for an understanding of skill is the nature of the spatial temporal organization of receptor effector feedback processes, which often are relatively independent of the specific receptor or effector elements initially involved

One of the clearest expositions of the view of skill to which most recent workers in the field subscribe was provided by Lashley (1951) in his Hixon Symposium lecture on "The Problem of Serial Order in Behavior," where he dealt with the "orderly arrangement of thought and action " Advocating the commonality of different forms of learning and learned behavior he argued (1951, p 113) that

> Certainly language presents in a most striking form the integrative functions that are characteristic of the cerebral cortex and that reach their highest developments in human thought processes Temporal integration is not found exclusively in language the coordination of leg movements in insects the song of birds the control of trotting and pacing in a gated horse, the rat running a maze the architect designing a house, and the carpenter sawing a board present a problem of sequences of action.

Later, in discussing the generality of the problem of temporal integration, which he referred to as the 'syntax of action," he argued that "not only speech but all skilled acts seem to involve the same problems of serial ordering, even down to the temporal coordination of muscular contractions in such a movement as reaching and grasping" (1951, p 123)

Spatial temporal patterning, the interplay of receptor-effector feedback

processes, and such characteristics as timing, anticipation, and the graded response are thus seen as indentifying characteristics of skill Examples of such behavior may be found on every hand Some forms of skilled behavior involve gross bodily activities, such as walking, running, jumping, dancing, swimming, and balancing Other forms of skilled behavior involve segments of the total response mechanism, as in reaching, grasping, and manipulating A great many skilled activities in the life of typical human beings today involve the manipulation of tools and objects or the control of machines, as is the case in writing typing, playing a musical instruments, sewing, driving a car, piloting an aircraft, playing tennis, throwing a ball, doing assembly work, or operating a rotary pursuit apparatus in a psychological laboratory

The use of language, under this definition, clearly is a form of highly skilled behavior However, in the interest of restricting the present topic to manageable proportions, language learning will not be considered as a primary problem, except insofar as it is necessary to redefine the topic of perceptual motor skill so that its relation to verbal, ideational, problem-solving, and information storage (memory) processes and to conditioning, paired associate learning, etc , is clarified

Taxonomy

Elements of skilled performance —Emphasis on organization and patterning of a behavior process immediately confronts the theorists, and the researcher, with the question of how the processes under study are to be analyzed I assume that some form of analysis is a necessary part of scientific study Much of the present paper either deals directly with this question of the basis for analysis of behavior sequences or discusses data from experiments that appear to have handled the issue in a satisfactory manner

Historically Robert S Woodworth (1899) is clearly the psychologist whose views on this question have had the most lasting influence on contemporary thinking, although Fullerton and Cattell (1892), Craik (1948), Stetson (1905), Bartlett (1958), and Montpellier (1937), among others, have made important contributions

In his later book on the *Dynamics of Behavior*, Woodworth (1958), extended his earlier (1899) view that the integration of behavior in the time domain can be understood in terms of "two-phase motor units," which usually occur in sequences or "polyphase motor units" under the integrating influence of "preset" and "retroflex," the latter being Troland's (1928) term for feedback

There is general agreement that the two-phase motor units are the building blocks out of which are fashioned spatially and temporally organ-

ized motor sequences The act of jumping may be cited as a typical example of a unified sequence consisting of a preparatory act (crouching) followed immediately by the act itself Hitting a golf ball is another example, here the backswing is the first part of the two phase unit The total time interval consumed by such a behavior unit is usually short, often about half a second Polyphase motor units, in simplest form, consist of sequences of two phase units Walking and rotary pursuit performance are good examples Although these ideas are found in Woodworth's 1958 book, they are also clearly present in his 1899 paper on the accuracy of voluntary movement In the earlier paper, for example, he made explicit the idea, subsequently accepted by almost all workers in the field, that the pattern for a typical two-phase motor response is set up in advance of its initiation, and may be uninfluenced during its execution either by sensory feedback, or by new independent sensory inputs He also concluded that in very rapid polyphasic responses, such as tapping successive targets, more than one two-phase response may be pre formed and emitted as a unit In passing, and as preparation for later reference to the analogy between skill learning and the development of a specific program for a data processing system, it is interesting to note that the setting up of a two phase motor response, in advance of its initiation, corresponds in many ways to the calling up of a computer *subroutine*, and that a repetitive polyphase unit corresponds in computer phraseology to what is called a *loop*

With the two-phase movement as a starting point, several additional taxonomic distinctions can now be made Although we are concerned specifically with perceptual motor skills, many of the task distinctions are equally valid for other types of tasks

Task continuity —One of the most important general characteristics of a task is its spatial and temporal continuity Since input, output, and feed back may each vary independently with regard to continuity, many com binations are possible The magnitude of temporal or spatial discontinui ties may also be important Thus if discontinuities are so small as not to be discriminable, then in effect the task is continuous Conversely, if con tinuous tasks contain marked periodicities, then they may become discrete from the S's viewpoint.

The source of pacing may also be an important task characteristic In either serial or continuous tasks the *S* may set his own rate (self pacing), as in typing or driving a car Or the task may be externally paced, as in taking dictation or in aiming at a moving target

Task coherence —The emphasis on spatial temporal patterning of behavior immediately suggests the importance of an adequate taxonomy for specifying the sequential organization of tasks The general term *coherence*

will be used to identify this task dimensinn, and tasks will be considered to differ in degree of coherence The twn most commonly used quaotutative indices of coherence include degree of *relative redundancy,* and degree of *autocorrelation* As an illustration, the rotary pursuit task is 100% coherent or redundant (if one assumes a perfect time clock or time sense) once frequency, pbase, and amplitude are specified Normal language behavior is variously estimated at arnund 50% redundant Tracking and information-handling tasks can be made to vary in coherence over wide limits, depending on the choice of input sequence by the experimenter, although it is seldom if ever possible in practice to produce a completely incoherent continuous signal (nne which is unlimited in frequency) It should be noted in passing that tbere is a trend in many areas of learning research to substitute the concept of coherence for that of meaningfulness Thus word association norms, and measures of the strength of population stereotypes in perceptual motor tasks can be viewed as defining the degree of coherence or stereotypy in the behaviors which *S*s bring to our experiments It is as important to specify the cobereoce of response sequences as of stimulus patterns

Complexity—Related to coherence, but capable of independent definition and manipulation, is tbe complexity of a task In simplest terms, complexity refers to the number of different stimuli, respooses or transformation operations that are possible (in a statistical or probabilistic sense) or that are actually cootained in some block of space or time (in a deterministic sense) Thus, the English alphabet is less conplex than the Chinese alphabet, and the 2 unit binary alphabet of a digital computer is even less complex Similarly, the instrument panel of an automobile is less complex than that of an aircraft

A *sequence* of binary symbols nr a sequeoce of alpba numeric symbols may be either highly coherent or incoherent The temporal intervals between significant events (nr the component frequencies of a continuous signal) as well as the events themselves, may be coherent or incoherent, and the events may be closely or widely spaced in time

Although there are important task dimensions in addition to continuity, coherence, and complexity, these three are sufficient to specify many of the most important general characteristics of perceptual motor tasks It should be re emphasized however, that the defining operations for each of these dimensions are applicable to aoy sequence of stimuli or responses

Theoretical Models

It is not possible, in this paper, to present a detailed framework for the study of skill learning In particular, it is not possible to discuss adequately

the three types of models which are stimulating much of the contemporary work in this area Nevertheless, a brief reference to each is necessary I shall assume some preliminary acquaintance with the models, and restrict my remarks to an emphasis on a few general concepts derived from each, especially those regarding control and communication processes which are most relevant for learning theory

Communication Models

Rather than viewing perceptual motor behavior as a series of motor responses made to reach some goal, it is possible, and I believe considerably more profitable, to view such behavior as an information processing activity guided by some general plan or program Information and communication concepts are now being applied to many kinds of processes and can easily be extended to models of skilled performance Information measures are useful in quantifying the processes involved in skills, but in themselves do not suggest any theory about human behavior or indicate the nature of the models that apply to human skills Analysis of the operation of an information flow or data processing *system*, however, has suggested several concepts that are beginning to exert a strong influence on psychological theory

The concept of *information processing* is one such general concept Thus skilled perceptual motor performance can be viewed as involving operations such as information translation, information transmission, information reduction, information collation, and in some cases the generation of information (or of noise) Information storage is also involved, of course

Central to all information processing is the general concept of *coding* A code consists of an alphabet plus a system of fixed constraints One goal of an informational analysis of behavior is to specify the codes involved in human behavior, including neurophysiological processes as well as the inputs to man s sensory channels and his response codes

It should be noted in passing that information codes may employ discrete or continuously variable signals However, most of the uses made of information measures and communication theory by psychologists have involved the assumption that man uses discrete codes i e , that he categorizes information

Emphasis on information processing and coding is one reason why the dichotomy between verbal and motor processes, such as between verbal and motor learning does not appear to be as important as it once seemed to be Man uses a very large number of information codes, and this in itself is an important topic However, the informational characteristics of any task can be considered independently of the particular code employed.

Control System Models

The feedback control system or servomechanism came into widespread use during World War II, although regulators and similar devices had been known long before At about the same time, work on skill learning received a strong impetus as a result of military interest in selecting and training aircraft pilots, gunners, and operators of other complex machines It is more than a coincidence that the pilot s job in many respects, corresponds to that of an autopilot, and that the variables which are considered by the design engineer in perfecting an automatic control system are analogous in many respects to the task variables which affect the learning of a pilot Several very general concepts of significance for psychological theory have been borrowed from control systems, of which the autopilot is one specific example F V Taylor (1957) called such very general models metaconcepts, referring to their usefulness in bridging the interface between physical and psychological science They should be equally useful in revealing commonalities in the field of human learning

The most generally used control system concept is that of *feedback* As noted previously, the idea of feedback (called by such names as retroflex and backlash) is an old one in psychology The concept of reinforcement in learning theory also emphasizes a form of feedback But the use of the feedback concept in a precise sense is relatively recent In particular, the precise specification of the nature of the feedback function and a distinction between input, feedback, and disturbance or noise is necessary before feedback theory can be made quantitative In this connection, it is important to remember that an important class of control system processes are those performed by regulators, in which the input is a fixed reference quantity, rather than a variable The maintenance of erect posture is such a regulatory process Here, all of the stimulation to which a person is responding is either feedback or some form of disturbance Driving a car on a straight road is also a regulatory process, the nearest we can come to identifying the inputs to a driver in such a situation is to specify the rules or instructions he has been given, plus the inferences we make as to his goals or purposes Otherwise, he responds entirely to feedback which tells him about his previous errors

Another useful control system concept is that of a mathematical representation of the relation of output to input, called a *transfer function* Given such a mathematical specification of the dynamics of a system, control theory can be applied by the engineer in order to predict performance or to modify the system so as to optimize some desired property Most problems in control system optimization center around questions of how

to modify the system transfer function directly, if this is possible, or if not, how to adjust feedback so as to optimize system performance With respect to feedback it should be kept in mind also that many different functions of the output may be fed back for comparison with the input Thus feedback may consist of integrals or derivations of the output, or samples of these or other output values As a corollary it appears that the human operator of a dynamic system must also learn the nature of the system dynamics and discover what behavior on his part will optimize system output This is true regardless of whether the task is a very simple one like throwing a rock or a complex one like controlling an aircraft Thus a rock has mass, the air offers resistance to its flight, and the rock-thrower must adapt to these simple ballistic characteristics An aircraft, of course, has much more complex characteristics, not the least of which is that its responses to control inputs vary as a function of altitude, airspeed, weight of fuel aboard, and many other variables

As is the case with an information system, a control system may be continuous or discrete However, the continuous system has received greatest emphasis Thus, the two models we have discussed are sometimes referred to as digital and analog systems, respectively Of the two, the information system is probably the more general model, since any control task may also be viewed as a special case of an information-processing task

Adaptive System Models

Since learning is an adaptive process, the usefulness of the two preceding models is limited by the fact that they are static, i e , they do not change their characteristics as a function of experience This limitation is overcome in large part by the adaptive system model, especially adaptive systems with memory (information storage) capacity

Several types of adaptive systems are currently of interest to physical and biological scientists One such group of scientists is studying adaptive feedback systems Others are studying artificial intelligence, growing automata and the like Still another group of investigators is studying the processes whereby stored program computers can be made to modify their own programs Further extensions of the adaptive system concept will undoubtedly appear soon (see Gibson, 1960)

Basic to the adaptive system is the existence of hierarchical processes Programs are provided for carrying out basic or routine functions, and other, higher-level programs or plans (see Miller, Galanter, & Pribram, 1960) are provided for modifying these lower-order ones on the basis of accumulated information Also intrinsic to the successful operation of an

adaptive system are three specific kinds of processes (*a*) one that insures variability in input or system parameters, (*b*) one that provides a criterion measure, and (*c*) one that results in the system changing or maintaining its program or parameters so that over time it will tend to achieve a performance level which is closer to the optimal One of the simplest such mechanisms is one that constantly searches for an optimum, but has a short memory so that it makes no use of previous search procedures and attacks each new problem in the same way The stored program computer does not have this limitation Newell, Shaw, and Simon have pointed out that

> The real importance of the digital computer for the theory of higher mental processes lies not merely in allowing us to realize such processes "in the metal" and outside the brain, but in providing us with a much profounder idea than we have hitherto had of the characteristics a mechanism must possess if it is to carry out complex information processing tasks' (1958, p 163)

The same writers also assert that many of "the vaguenesses that have plagued the theory of higher mental processes and other parts of psychology disappear when the phenomena are described as programs" (1958, p 166) The concept of behavior organization in skilled activities also becomes much clearer and more operational when defined in such terms

Adaptive processes may be continuous or discrete More commonly, lower-order processes may be continuous, while higher-order processes are discrete, i e , effect discrete changes in lower-order ones usually over a relatively long time cycle As an illustration, it was found in connection with early efforts to apply linear, static-model feedback theory to human tracking performance that over short periods of time and for limited task conditions human tracking behavior may be described adequately by means of a fixed linear model or simple transfer function but that over longer periods of time tracking behavior is likely to exhibit discontinuities and be highly nonlinear

For the student of learning, the most promising model of an adaptive process is that provided by the stored-program data processing system The subroutines of such a program may be modified, so as to become more efficient, or the higher level "executive" program, which calls up the various subroutines, may be improved However, the program itself is quite independent of the particular data on which it operates at a given time Such programs can be written for many different communication, control, and data processing tasks, varying from the processing of sensor inputs, to tracking and playing chess

A Composite Model

In summary, if one views perceptual motor skill as a composite of communication, control, and data-processing activities, then the perceptual-motor tasks that we ask *Ss* to learn can be specified by reference to the abstract properties of sequences of events or signals (such as their continuity, coherence, and complexity), to the dynamics of the system (as specified by its transfer function), to the nature of feedback functions, and to the subroutines and executive programs of an abstract data processing system Within such a general framework it is also possible to employ several additional important concepts such as information coding and information transformation and processing, and, more important, to develop specific models of adaptive control and information handling processes It remains to demonstrate how this general framework will aid us in understanding the nature of skill learning, and how general the framework is

The first theoretical issue to be considered within this framework will be the one touched upon earlier, when the question was raised, What is the basic unit or element of skilled performance? The fundamental issue is the suitability of a discrete as contrasted with a continuous model of skilled performance

The Continuity-Discontinuity Issue in Skill Learning

Learning theorists are familiar with one aspect of the continuity issue in learning This is the question Are learning changes gradual or sudden, continuous or discontinuous? An example is the notion of sudden or 'insightful' reorganization of behavior The study of skilled performance raises another aspect of the continuity-discontinuity issue, an aspect which has already come up repeatedly in the preceding discussion of continuous (or analog) vs discrete (or digital) models Here the issue concerns both the continuity of learning or adaptive processes, and the continuity of underlying perceptual motor processes This issue has interested many recent students of skilled performance

The issue is an important one because the answer dictates the kind of model that is appropriate, and the kinds of experimental tasks and methods that are most appropriate for the study of skill In particular, the generality of much current research on human information handling performance hinges on the answer to this issue In short, are two theories of skill learning and performance needed, one for continuous and one for discrete tasks, or will the discrete model suffice for both types of tasks? The sources of evidence on this issue are diverse, and none of the data are

highly conclusive We shall therefore review several lines of evidence briefly

Evidence on Continuity-Discontinuity

Limited capacity for discrimination, perception, and short-term memory —It is clear that most perceptual-motor tasks require discrimination (usually on an absolute basis), classification, and other perceptual processes, as well as short-term memory and response differentiation These processes are usually carried out under a certain amount of time stress Very significant, from a theoretical view, is the close agreement among estimates of the absolute capacities of individuals in accomplishing these different functions (see Miller, 1956, and Newman, 1959, for example) Capacity for absolute judgment of stimulus magnitudes, span of perception, short-term memory, capacity for the discrimination of proprioceptive feedback, and capacity for the reproduction of quick movements all seem to be very closely matched when viewed as information and control processes

Now this in itself is not conclusive evidence in favor either of the discrete or the continuous type of model However, it is entirely consistent with the concept of a system that handles information in discrete "chunks," that employs numerous steps which are accomplished in series, and that consists of matched components

Intrinsic periodicities in response patterns —One of the first writers to raise the continuity issue in relation to perceptual-motor skill was Craik (1948) who observed highly consistent discontinuities in graphic records taken in continuous tracking tasks He concluded that *S*s who track a constantly moving target often show irregularities in their responses which recur once or twice a second Such irregularities are much more marked early in learning than later, as shown in Fig 1, which is reproduced from Fitts, Bahrick, Noble, and Briggs (in press) Unfortunately, several types of artifacts arise in connection with the use of such graphic records First, direct visual analysis of graphic records is highly unreliable, so that it is necessary to use autocorrelation or frequency analysis Second, the presence of periodicities in the *S*'s error record which are of the same frequency as the input tends to obscure any periodicities introduced by *S* himself Such effects are clearly shown in the error autocorrelation records in Fig 2 (from Fitts, Bennett, & Bahrick, 1957). As practice is continued, less and less of the target periodicity comes through in the error record, but any additional periodicities introduced by *S* also become less conspicious A few efforts have been made to obtain autocorrelation records of error early in tracking tasks having zero or constant velocity inputs (thus avoiding the input frequency artifact), but this approach has not

been sufficiently exploited to provide us with conclusive answers At present, therefore, the data merely suggest that *S*s may, at least early in a continuous tracking task, make intermittent corrections at a rate up to about 2 per sec , but we cannot be sure of the effect

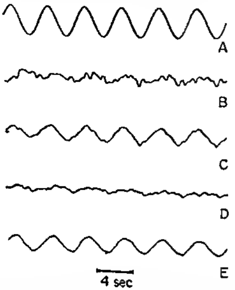

FIG 1 Periodicities in motor (tracking) responses to a slow (one cycle per 4 sec) sinusoidal target motion. Stimulus motion is shown in A C and E are the *S*'s responses early and later in practice, B and D are corresponding error records, early and later in practice Note that the irregularities introduced by the *S* appear to have a frequency of 1 to 2 per sec

Subjects are certainly able to learn to make fairly smooth, continuous movements, however These movement patterns may last for periods of at least several seconds or several cycles of a polyphasic response and can be carried out for a time with the eyes closed and no apparent exteroceptive feedback If we examine graphic records of such highly coherent responses carefully, however, especially as we push *S* to some limit in a visually controlled task, we may again note evidence of periodicity, such as periodic adjustments in the amplitude or frequency of responding every several cycles Such evidence is apparent, for example, in the "human frequency response" data shown in Fig 3 (from Noble, Fitts, & Warren, 1955), where *S*s were tracking a sinusoidal input of several cycles per second Woodworth (1899) observed similar periodic adjustments in *S*s' behavior in serial dotting tasks, and concluded that more than one biphasic movement may be pre selected and emitted as a unit

Reaction time and the psychological refractory period—One of the most commonly studied characteristics of perceptual motor performance

been sufficiently exploited to provide us with conclusive answers At present, therefore, the data merely suggest that *S*s may, at least early in a continuous tracking task, make intermittent corrections at a rate up to about 2 per sec, but we cannot be sure of the effect

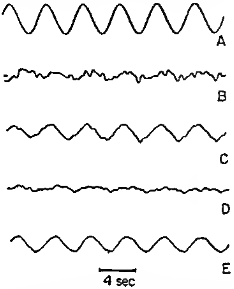

FIG 1 Periodicities in motor (tracking) responses to a slow (one cycle per 4 sec) sinusoidal target motion. Stimulus motion is shown in A C and E are the *S*'s responses early and later in practice B and D are corresponding error records, early and later in practice Note that the irregularities introduced by the *S* appear to have a frequency of 1 to 2 per sec

Subjects are certainly able to learn to make fairly smooth, continuous movements, however These movement patterns may last for periods of at least several seconds or several cycles of a polyphasic response and can be carried out for a time with the eyes closed and no apparent exteroceptive feedback If we examine graphic records of such highly coherent responses carefully, however, especially as we push *S* to some limit in a visually controlled task, we may again note evidence of periodicity, such as periodic adjustments in the amplitude or frequency of responding every several cycles Such evidence is apparent, for example, in the "human frequency response" data shown in Fig 3 (from Noble, Fitts, & Warren, 1955), where *S*s were tracking a sinusoidal input of several cycles per second Woodworth (1899) observed similar periodic adjustments in *S*s' behavior in serial dotting tasks, and concluded that more than one biphasic movement may be pre selected and emitted as a unit

Reaction time and the psychological refractory period—One of the most commonly studied characteristics of perceptual motor performance

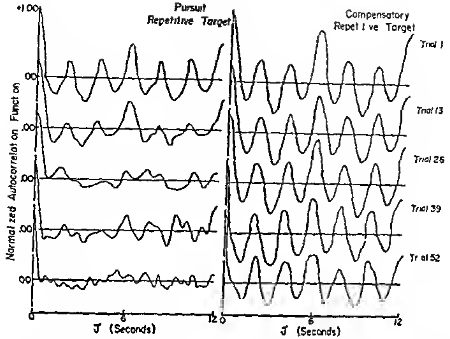

FIG 2 Autocorrelation records of manual tracking error records at different stages of practice, for a pursuit and a compensatory display (from Fitts, Bennett, & Bahrick, 1956)

is human reaction time in tasks where there is temporal or event uncertainty It is very difficult to measure reaction (lag) time in a continuous task But *it is easy to present two discrete stimuli in rapid succession* and to observe the effects of the second stimulus on the reaction time to the

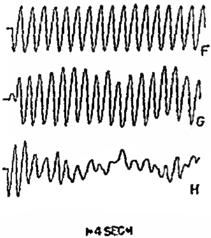

FIG 3 Periodicities in motor (tracking) responses to a fast (1 cycle/sec) sinusoidal target motion (F) The motor response of the *S* is shown in G and error in H Note the scalloping of the responses in G

first, and the effect of the first S–R event on the reaction time to the second stimulus At least it is easy in principle A substantial amount of research has been done on this problem The results clearly indicate that a preceding stimulus may delay the response to another stimulus which follows it if the interval between the two is between about 0 1 to 0 3 sec These results, however, have been interpreted both as an indication of the validity of the discrete model, and in terms of non-optimal set or expectancy

Hirsh and Sherrick (1961) report that two stimuli must be separated by at least 0 40 sec for their order of occurrence to be discriminated correctly 95% of the time Such data are congruent with the theory of a psychological refractory period, as advanced by Craik (1948), Hick (1948, Welford (1952), and others This theory suggests that the time for a single complete response cycle (discrimination of stimulus—response —discrimination of feedback) may be the basis for one kind of human intermittency

Information handling rate in continuous, serial and discrete tasks— Another line of evidence bearing on the continuity issue is the comparative rate of handling information in different types of tasks Most estimates of the upper limit of performance in speeded perceptual motor tasks have been around 25 to 35 bits per second for highly practiced *Ss* in activities such as speaking reading, piano playing, and typing (Newman, 1959, Pierce & Karlin 1957) Although the upper limit of information handling rate varies markedly with several variables such as learning and coding (Alluisi, 1957), no one has as yet proposed that the difference between discrete and serial tasks, per se, is an especially important variable in this connection Speaking, silent reading, continuous tracking, and serial key-pressing for example, all give roughly similar estimates of peak performance capacity

A study by Brainard, Irby, Fitts, and Alluisi (1962) which used both a serial task having a 2 sec delay between each response and the next stimulus and a typical discrete reaction time task employing a 2 sec warning signal and 10 sec between stimuli provides a direct comparison of serial and discrete performance The results (see Fig 4) indicate small although fairly consistent differences in the two tasks Errors were slightly fewer for the discrete case, but response times were slightly shorter for the serial case, so that when rate of information transmitted is used as a criterion, performance is very similar in each of the four S–R tasks studied

Some years ago (Fitts 1954), I published some data on information rate in controlling the amplitude of movement in continuous, cyclical tasks Subjects attempted to make alternate hits on two targets, as shown in

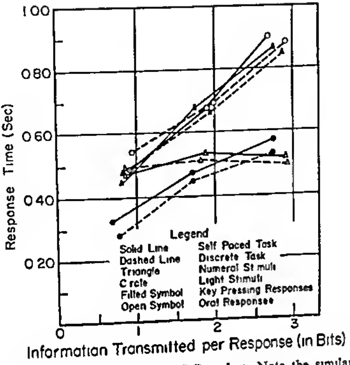

FIG 4 Reaction time as a function of S-R coding Note the similarity of times for self paced serial and for discrete tasks (after Brainard, Irby, Fitts & Alluisi, 1962)

Fig 5, working as rapidly as they could This can be classified as a continuous task (although hits were recorded discretely) since the two targets were continuously present and there were no enforced pauses or discontinuities Movement amplitude and required movement accuracy were varied systematically, and average time per movement cycle and errors were recorded Recently, we (Fitts & Peterson, in press) have completed a similar study using discrete responses in a typical 2 choice reaction time experiment The apparatus is shown in Fig 6 Subjects were sometimes required to hit one target as soon as possible after a light came on (knowing in advance which light-target pair would be used), at other times they were told to hit one of two targets as soon as one of two possible lights came on A 2-sec warning signal was used in both tasks, in the 1-choice task, the one stimulus light appeared with a probability of 0 5 after the warning Under these task conditions, reaction time was found to change very slightly as a function of the relationship between movement amplitude and accuracy The latter results for the two studies using continuous and discrete tasks are shown in Fig 7 In both studies the average time taken to execute a movement increased linearly as a function of the amount of information which the movement was required to

generate However, the continuous task was considerably less efficient, i e, comparable movements took considerably longer than they did in the discrete task A likely explanation for this difference is that in the continuous task, Ss had to insert a reaction time every few cycles of the movement in order to evaluate feedback data and keep the process under control This idea, of course, is consistent with earlier notions of Woodworth, Welford, and others regarding human intermittency Furthermore time in contact with the targets was included in the overall time for the continuous task, but was not included in movement time for the discrete case

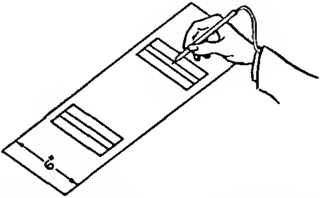

FIG 5 Apparatus used in studies of the effect of movement amplitude (A) and target width (required response accuracy W) on rate of responding in a continuous task (after Fitts 1954)

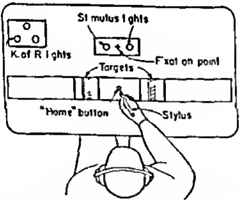

FIG 6 Apparatus used in studies of the effect of movement amplitude (A) and target width (W) on reaction time and movement time in a discrete two-alternative task (after Fitts & Peterson, in press)

Finally the data on eye fixations in reading and similar visual tasks should be mentioned The motor system of the eye is very efficient but its maximum rate is about five saccadic fixation movements per second,

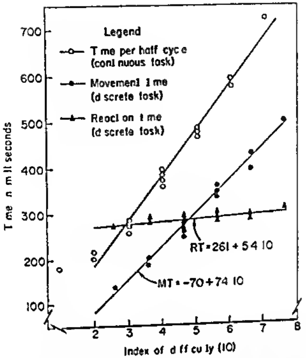

FIG 7 Reaction time and movement time for discrete responses as a function of the index of movement difficulty (ID) (after Fitts and Peterson 1964) The top curve (open circles) are comparable cycle time data for a continuous task (Fitts 1954)

and in many perceptual motor tasks, such as piloting an aircraft the rate is only about two fixations per second Visual input under such circumstances is clearly discrete, and since estimates of immediate memory capacity are that the *S* can hold only the information obtained in one or two fixations, the eye seems to be well adapted to use by a system which is operating discretely on one or two eye fixations worth of data per response cycle

Summary and Implications

Although the evidence is quite diffuse and none of it is very conclusive with respect to the continuity discontinuity issue, it does appear that a discrete model is adequate for describing behavior in continuous as well as in discrete tasks Viewing skilled performance as an information processing task, but making use of concepts borrowed from feedback and adaptive system theory, a model for perceptual motor skill learning might take the following general form (for a more complete discussion see Fitts 1962)

An adult or even a child of a few years of age, never begins the acquisition of a new form of skilled behavior except from the background of many already existing highly developed, both general and specific

skills Thus the initial state of our model is not that of a random network, but an already highly organized system possessing language skills, concepts, and many efficient subroutines such as those employed in maintaining posture, walking and manipulating The number of such identifiable highly developed skills in an adult is certainly in the hundreds, each having its own executive program and library of subroutines, many of the subroutines being shared with other skills

Learning to swim provides a typical example of a skill that is learned against a complex background of already existing habits The first (hypothetical) step in such learning is the setting up of a general executive program What usually happens in such a learning situation is that *S* listens to instructions, observes demonstrations, and tries out different routines which he already has available, until somehow or other he gets started at the learning task Verbal mediation plays an important role in this early stage

The actual sequence of behavior processes employed early in learning varies with the type of activity, of course, but might be somewhat as follows The *S* observes or samples certain aspects of the environment, puts this information in short term storage after some recoding, makes a decision such as selecting an appropriate subroutine which sets up a response pattern, executes a short behavior sequence such as a biphasic or polyphasic movement, samples the internal and external feedback from this response plus additional stimulus information from the environment, recodes and stores this new information (in the process losing some of the information already in short term storage), makes another decision which might be to use a different subroutine, and so on As learning progresses, the subroutines become longer, the executive routine or overall strategy is perfected, the stimulus sampling becomes less frequent and the coding more efficient, and different aspects of the activity become integrated or coordinated (such as kicking, breathing, and use of the arms in swimming) In other types of perceptual-motor tasks, such as those which are less coherent than swimming, the improvement may take the form of strategies and decisions processes better adapted to the probabilities associated with stimulus sequences As learning continues, overall performance may come to resemble more and more closely a continuous process The overall program having now been perfected, frequent changes no longer need to be made in it However, subroutines may continue slowly to become more efficient, and the *S* to become increasingly able to carry on the entire behavior process while engaged simultaneously in other activities, with little or no interference between the two

Such a general verbal description of skilled learning probably sounds quite familiar to anyone who has read the older literature on this topic

Consider, for example, the following brief quotation from Bryan and Harter (1899, p. 373) on the learning of telegraphy:

> A plateau in the curve means that the lower order habits are approaching their maximum development but are not yet sufficiently automatic to leave the attention free to attack the higher order habits.

One of the most important advances made in the years since Bryan and Harter's studies of telegraphy is clarification of what is meant by a program or plan (see Miller, Galanter, & Pribram, 1960) governing a sequence of operations. The present writer has been working for some time on a computer program to simulate a hypothetical batter hitting a baseball, the batter having available sensors for sampling the flight of the ball, a computer for determining azimuth and trajectory, a memory for baseball "lore," a probability computer and the like, but operating in real (human) time with an assumed cycle time corresponding to a human baseball player. This does not appear to be an appropriate time to unmask this particular automaton, but my own conclusion is that if a digital computer can be programmed to play chess against a human opponent, then it can probably be programmed to hit a baseball thrown by a good human pitcher, a skill not to be dismissed lightly, since in many respects it is much more complex than that involved in playing chess.

In the remainder of this paper, skill learning will be treated as if the discrete model were the appropriate one. I assume that my colleagues who like to enumerate stimuli and responses, to refer to the number of discrete reinforcements, and the like, will be pleased by this conclusion and will find it acceptable if I now sometimes talk about such things as rate of responding and frequency of reinforcement in a continuous tracking task. Specifically, after a decade of research devoted chiefly to the study of skilled performance in continuous tasks, I have recently turned to the study of information handling behavior in serial and discrete tasks, but I believe that I am studying essentially the same basic perceptual motor skill processes as before. Fortunately, as mentioned earlier, all of the task taxonomy which I shall employ is equally applicable to continuous and to discrete sequences of behavior.

Phases Characteristic of Skill Learning

Skill learning is primarily a continuous process even though the fine grain structure of the performance itself may involve discrete operations. Thus it is misleading to assume distinct stages in skill learning. Instead, we should think of gradual shifts in the factor structure of skills, or in the nature of the processes (strategies and tactics, executive routines and subroutines) employed, as learning progresses. The evolving process is

revealed by the organization of behavior into larger and larger units, as Bryan and Harter (1899) emphasized, and toward hierarchical organization, as Miller, Galanter, and Pribram (1960) have recently emphasized in their discussion of motor skills and habits

Changing Factor Structure of Skills

Recent correlational analyses of performance at different points in skill learning reveal significant changes in the relationships among abilities both in different tasks and at different stages of practice in the same task Correlations between the first trial and successively remote trials become progressively lower, whereas correlations between the most recent adjacent trials become progressively higher Also, the factor structures of complex tasks change consistently with practice, indicating that ability requirements are different at different stages of learning (Bilodeau & Bilodeau, 1961, Fleishman & Hempel, 1954, 1955), as Bryan and Harter suggested long ago

Phases in Skill Learning

Early phase—The earliest phase of skill learning in an adult may be of very short duration in simple tasks, covering only the time required to understand instructions, to complete a few preliminary trials, and to establish the proper cognitive set for the task

Skill learning processes during this phase are undoubtedly very similar to those involved in the early phases of rote learning Underwood and Schulz s (1960) discussion of the response learning stage of rote learning applies equally well to skill learning, especially when the responses are heterogeneous and the total situation is new Response integration is especially important when the new task requires the simultaneous use of two previously differentiated sets of responses, or response subroutines For example, response response compatibility effects are important determinants of difficulty in the learning of complex tasks involving both bands or hands and feet, as when a typist first tries to use the foot pedal to govern the play back of a dictating machine when a beginning student of music first tries to produce different rhythms with the two hands, or when a novice first tries to coordinate breathing and arm strokes in swimming

Intermediate phase—The intermediate phase of skill learning resembles Underwood s ' hook up or associative stage of rote learning Two kinds of mediation processes appear to be important at this time One mediates the formation of specific associations and learning to respond to specific cues The other involves cognitive set learning discussion of which will be deferred briefly

An unpublished experiment completed recently by Fitts and Switzer

illustrates clearly the role of mediating associations in learning and also shows the close relationship between perceptual-motor and verbal learning The task was to make a vocal response as soon as possible after the exposure of a picture Twelve pictures were selected to represent objects whose names were of highest frequency in the Thorndike-Lorge word list, with the restriction that each name begin with a different letter of the alphabet The pictures are shown in Fig 8 They were exposed by the opening of a double-bladed mechanical shutter Vocal reaction time was detected by use of a boom microphone and voice key and measured to the nearest 01 sec The effects of two variables were studied—number of alternative stimuli (12 vs 3 pictures), and directness of the verbal mediation All responses involved saying a letter of/the alphabet In one case (direct mediation), the response was the first letter in the familiar name of the object In the other case (indirect mediation), the response was one of the other 11 letters (selected randomly) Note that neither response is one that *S* would ordinarily make The population stereotype is to give the name of the object not a letter of the alphabet In one case it is assumed that the required response was "hooked up" to the stimulus, by an associative chain consisting of Stimulus → Familiar object name → Vocalization of first letter of name In the other case, presumably the associative chain involved at least one additional step Stimulus → Familiar object name → Some other name → Vocalization of the first letter of the other name When interrogated at the end of the experiment, most *Ss* reported such a chain of associations Subjects were pre-trained to a criterion of two correct trials by an efficient training method, and then reaction times were taken during five 30 min testing sessions Reaction time data are

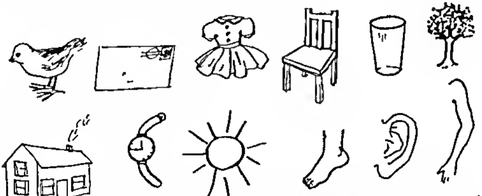

FIG 8 Pictures of common objects used in an unpublished study by Fitts and Switzer The responses to these pictures which were assumed to be mediated by "familiar associations were bottom row, left to right, H, W, S, F, E, A Top row B, L, D, C, G, T Unfamiliar associations employed the same set of letters but different pairings

are each for a single *S* and thus are free of the artifacts that frequently appear in group data. Each of these studies, by a different investigator, indicates that the log-log relation between time and trials is essentially linear after the first few thousand trials.

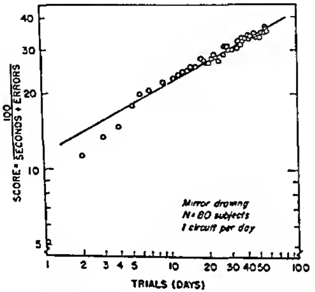

FIG. 10. Gradual improvement in mirror drawing with long-continued practice (after Snoddy, 1926).

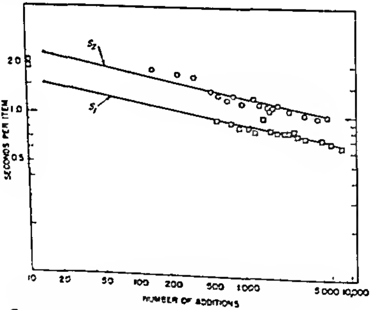

FIG. 11. Gradual improvement in mental arithmetic with long-continued practice (after Blackburn, 1936).

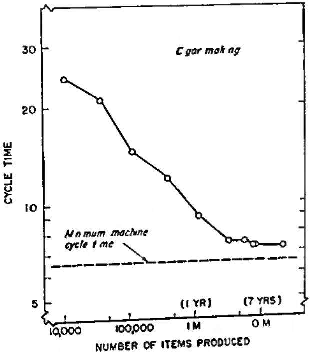

FIG 12 Gradual improvement in the performance of an industrial task over several years of work (after Crossman 1959)

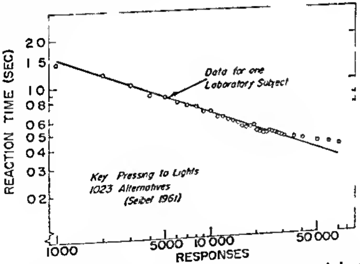

FIG 13 Gradual improvement in a 10 bit key pressing task with long-continued practice (after Seibel as reported in Klemmer 1962)

Another line of evidence regarding the continued improvement of performance in perceptual motor skills comes from case studies of the conditions of practice necessary for attainment of championship performance in individual athletic competitions, such as skating swimming diving and

track events, in games of skill such as bowling or golf, in competitive sports such as baseball and football, and in artistic performances such as singing and playing musical instruments It is very rare for peak performance in any of these activities to be reached short of several years of intensive, almost daily practice And the fact that performance ever levels off at all appears to be due as much to the effects of physiological aging and/or loss of motivation as to the reaching of a true learning asymptote or limit in capacity for further improvement Thus, in the case of skill learning, the asymptote, along with the plateau, must be viewed as an exception, rather than an accepted phenomenon of learning

Two kinds of evidence from developmental studies also indicate that perceptual motor behavior develops slowly on the basis of a great deal of practice Restriction of early visual motor behavior (in the case of animals) provides one kind of evidence Riesen and Aarons (1959), for example, find that animals whose only early visual experience is gained with the head and body immobilized subsequently have great difficulty learning to control their own locomotion by the aid of vision The other line of evidence comes from studies in which the natural relation between visual cues and motor behavior is disturbed, (Snyder & Pronko, 1952, Ruhle & Smith, 1959) or in which *S*s are asked to perform visual motor tasks to which the required responses are contrary to cultural patterns and population stereotypes (Fitts & Seeger, 1953) Although subjective reports often indicate that the old visual movement habits may be unlearned and replaced by new ones after a few days, precise measures of skilled performance show that decrements persist over the maximum periods of time yet studied (see Fig 14)

Cognitive Aspects of Skill Learning

Specific S–R Associations vs Cognitive Set Effects

The issue of specific versus generalized learning effects, i e , the learning and use of specific associations versus the learning and use of generalized sets or concepts, is an old one in learning theory and has its counterpart at all stages of skill learning The issue seems to be largely one of relative importance In the case of perceptual motor skills the importance of cognitive or mediational processes relative to simpler association processes appears to depend primarily on one of the task variables mentioned earlier —task coherence In a highly coherent task, regardless of the length or complexity of the sequence, stimulus and response patterns are essentially fixed and it would seem that the middle and late stages of the learning of such fixed response sequences should reflect a process similar to that governing the associative phase of rote learning and conditioning In the case

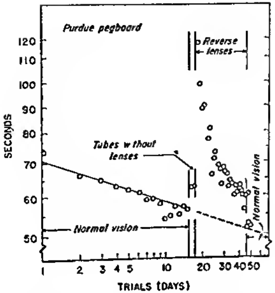

FIG 14 Effect of the wearing of reversing lenses on a task requiring precise visual control of motor responses (after Snyder & Pronko 1952) Note that at no time did performance while wearing the reversing lenses reach the earlier level achieved with normal vision or attained one day after their removal

of less coherent tasks, however, very large numbers of response patterns are involved and probabilistic rather than deterministic rules govern events Here it seems reasonable that strategies and executive routines become increasingly important, probability densities have to be learned rather than discrete events, and the generalization of learning is mediated increasingly by cognitive sets

The learning of language and perceptual motor skills frequently requires the development and use of very general cognitive sets As an illustration, an individual often responds to the same word very differently depending on the context He adjusts quickly to instructions such as "now respond with a word having the *opposite* meaning," or "tell me the *class* to which the following word belongs" Similar instances of the importance of set are observed in the realm of perceptual motor behavior Lashley used the term "syntax of action" to emphasize this relation For example, the response of an aircraft pilot to a statement such as "right wing down" (which can mean two opposite things) will depend on whether the statement is interpreted as a command, or as a report of error Similarly the movement of a needle on a dial can also be interpreted in two opposing ways, and the set of the observer will often determine whether he moves the related control in one direction or the other

The point to emphasize here is not so much that people develop cognitive or learning sets, or show adaptation level phenomena in perceptual-

motor tasks, but that they can develop many different cognitive sets, can switch from one to another readily, and can include the same stimulus or response elements as members of many different cognitive sets

Cognitive Set Learning in Skills

Viewing perceptual motor performance as an information handling skill suggests several ways of studying cognitive set learning its role in skilled performance, and its relation to stimulus and task variables Several of these topics will now be considered.

Compatibility effects and number of alternatives—S–R compatibility effects are ordinarily defined (Fitts & Deininger, 1954) in terms of performance or learning changes attributable to the interaction (congruence) of stimulus and response sets[2] They can easily be demonstrated in experiments where two or more sets of stimuli are paired with two or more sets of responses Recently it has become apparent that the absolute magnitude of such effects tends to be greater, the greater the complexity of the task (e g , the greater the uncertainty per stimulus) Griew (1958) has pointed this out, and some recent unpublished experiments by Fitts and Peterson show it clearly In Fig 15, for example, are shown three functions relating choice reaction time to average amount of information transmitted per response The bottom curve is for the most highly compatible task we have studied—pointing one's finger at a light The *S* simply moves his finger from a starting position and touches whichever one of a set of *n* lights is presented by *E* In this task reaction time increased by only 017 sec for each 1-bit increase in stimulus uncertainty A second function was determined for the same pointing responses, but with the hand and the response targets hidden from view by a screen and with the lights which served as stimuli located on a vertical panel in front of *S* Reaction times are still quite rapid but slightly slower than before, and the effect was greater for 9 than for 3 alternatives The top curve in Fig 15 is included for comparative purposes it is based on Hick s (1952) data where the task was to push a finger key in response to a light The slope for Hick's data is 11 sec per bit Even steeper slopes are found when finger responses must be made to numerals rather than to lights Hyman (1953) reported a slope of 18 sec per bit for such a task and the data of Brainard et al (1962) presented earlier also give 18 sec per bit Thus the slope of these last

[2] The term *set* as used in this paper has two meanings Cognitive set refers to preparation in advance for the probabilities or contingencies characterizing a given situation This meaning of the term set is similar to what Miller Galanter and Pribram (1960) call a Plan. The other usage of the word set is the mathematical one meaning a number of things of the same kind. Fortunately the two meanings are congruent.

functions are 10 times as steep as for the pointing responses shown in Fig 15

These data are subject to an interesting interpretation from the view

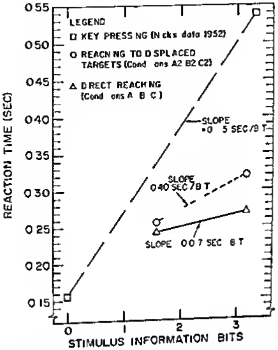

FIG 15 Effects of increasing stimulus uncertainty on choice reaction time for tasks having different levels of S–R compatibility The upper curve is from Hick (1952) for key pressing to lights the lower curves are data from Fitts & Peterson (in press) for highly compatible pointing responses to lights

point of learning theory As we progress from tasks low in compatibility to ones of relatively high compatibility, Ss are presumably making more and more use of very well established habits (i e, using responses which show strong population stereotypes) It appears, therefore, that the effects of stimulus uncertainty gradually become less and less marked in magnitude as a result of continued practice in information handling tasks There is considerable evidence, not reviewed here, for this conclusion On the surface this empirical finding can easily be interpreted as supporting the notion of the learning of specific S–R associations However, other experiments, which will now be described, demonstrate that part of the effect is due to cognitive set learning (i e, learning which has to do with sets of stimuli or sets of responses)

Subset familiarity—In studies of learning in an information handling task as a function of stimulus uncertainty one is confronted with a problem very similar to that involved in attempting to construct different sets of

materials equated for meaningfulness for use in a rote learning experiment This is the problem of controlling the amount of previous experience with alphabets, or S–R ensembles, of different size In some instances the stimulus side of the problem can be solved by using lights, pictures or other types of symbols which are drawn from sets of indeterminate maximum size Subsets of symbols such as numerals and letters, drawn from alphabets of a known and fixed size, however, are immediately suspect To put it simply, *S*s may continue to respond as if the entire alphabet were possible, even though the experimenter uses only a subset of the available symbols

Fitts and Switzer (1962) have recently demonstrated such effects in a series of three experiments In one experiment three sets of numerals and three groups of *S*s were used The task was very simple—to say the name of a numeral as soon as it was exposed One group worked with eight numerals, 1 through 8 A second group used an unfamiliar subset of two numerals, 2 and 7 A third group used a relatively familiar subset of two numerals, 1 and 2 Results, for three training sessions, are shown in Fig 16 Only the data for the numeral 2, which is common to all three groups, are shown On the first session vocal reaction time was the same for the numeral 2 when it was one of eight numerals and when it was a member of an unfamiliar subset of two numerals When the same numeral appeared as one of a familiar subset of two stimuli, however, reaction time was faster Not unexpectedly, learning was fastest for the small, unfamiliar subset Similar results were found in two other experiments with alphabetic symbols Figure 17 summarizes results for the three groups in this study which used, respectively, all 26 letters, an unfamiliar subset (EBP) and a familiar

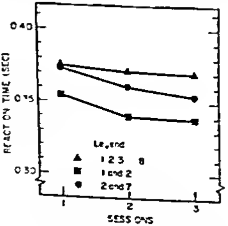

Fig 16 Choice reaction time in vocalizing the name of the numeral 2, as a function of the set of alternative numerals with which 2 is associated (after Fitts & Switzer 1962)

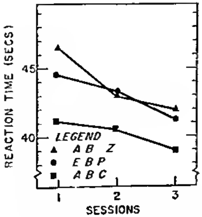

FIG 17 Choice reaction time in vocalizing the name of the letter B as a function of the set of alternative letters with which B is associated (after Fitts & Switzer, 1962)

subset (ABC) The data shown here are for the letter B which was common to all three sets The results are similar to those for the previous experiment and are significant statistically However, it should be noted that in all of these experiments the absolute differences are small, in agreement with the general finding for other highly compatible tasks

These results show clearly that even in the case of S–R associations that presumably have been practiced almost daily for at least fifteen years, cognitive sets are operating to some extent This offers support to the view that cognitive factors are important in all types of highly practiced skills The importance of cognitive set effects is further indicated by experiments with redundant sequences

Redundant sequences—We turn now to a topic which is at the heart of perceptual motor skill learning This is learning about sequences or patterns of events, where the patterns are probabilistic rather than deterministic

James McKeen Cattell (1886) was the first to study this problem from what is now called an informational point of view, working in this country and in Wundt's laboratory in Germany Cattell made up sequences of letters and of words which the *Ss* (Cattell himself and such volunteers as G O Berger, John Dewey, and G Stanley Hall) read serially as rapidly as possible In some instances letter and word sequences were those of English sentences, in other instances letters and words from English text were printed in reverse order thus giving unfamiliar sequences Cattell found that " it takes about twice as long to read (aloud, as fast as possible) words which have no connexion as words which make sentences, and letters which have no connexion as letters which make words" (1886, p 64) He then verified these results with text taken from French, Ger-

man, Italian, Latin, and Greek and found that the magnitude of the effect attributable to language structure was proportional to familiarity with the language Cattell's discrete vocal reaction times, incidentally, were of about the same absolute magnitude as those shown in Fig 17 (slightly less than 0 5 sec.).

Fitts, Wolpe, and Peterson (1963) have recently completed several studies in which they have measured discrete vocal reaction times to the elements of redundant sequences of numerals Holding the size of the set of numerals fixed at nine (1 through 9), one of the numerals was made more and more probable, all the remainder being made equally and increasingly less probable Maximum redundancy involved a sequence of 126 stimuli out of which 118 were the frequent symbol and each of the other eight symbols appeared only once ($R = 87\%$). Reaction time data for different groups of *Ss* are shown in Fig 18 All numerals were equally probable on Session 1; on the next three sessions each group worked at a different level of redundancy Differences in reaction times to frequent vs infrequent elements increased as a function of degree of redundancy. Unfortunately, as it turned out, the numeral 1, which was chosen as the most frequent numeral in the first experiment, gave slightly faster reaction times than the average for the remaining stimuli under the equal frequency

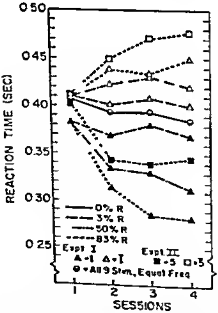

FIG 18 Differential choice reaction times in vocalizing the names of frequent and infrequent numerals as a function of increasing frequency unbalance (redundancy) (after Fitts, Peterson & Wolpe, 1963) In Exp 1 the numeral 1 was most frequent, in Exp 2 the numeral 5 was most frequent.

condition Therefore, one additional high redundancy group was run with the numeral 5 as the most frequent element The results for this group are very similar to those for the previous groups when the latter are corrected for the slight initial difference between 1's and other numerals These data, and some comparable results from a very different and much more compatible task are also shown in Fig 19 The difference in reaction times to infrequent vs frequent stimuli apparently increases linearly with redundancy, the slope of the function decreasing with degree of S–R compatibility

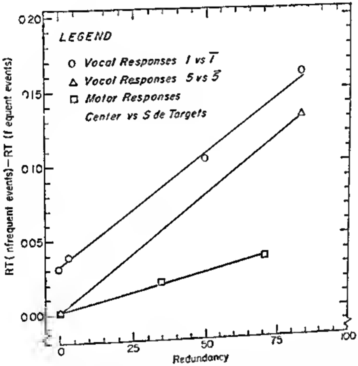

FIG 19 Differences in vocal reaction times to frequent vs infrequent elements of a set expressed as a function of the redundancy of the stimulus sequence The two top curves are for the last two sessions of the data shown in Fig 18 the bottom curve is for a task in which S had to make highly compatible pointing responses to lights (after Fitts Peterson & Wolpe 1963)

The error data from these reaction time experiments present another and highly important side of this picture Total errors remained highly stable at around 1% of all responses throughout the four sessions, but the proportions of errors made to the infrequent stimuli increased steadily relative to those made to the frequent stimuli

Space does not permit a full discussion of the theoretical significance of these error data, which can be interpreted in terms of a sequential stimulus

sampling theory It is obvious, however, that speed and accuracy in responding to any particular stimulus element in a sequence is influenced markedly by the probabilities governing its occurrence in the sequence This general finding is, of course, an old and well known one However, the demonstration of the extent and lawfulness of the effect, in terms of stimulus redundancy, is new, and of considerable theoretical importance This finding is strongly suggestive of operations analogous to a sequential statistical decision process Stated simply, a sequential statistical decision model of an information handling process hypothesizes that a cognitive set is first established which assumes certain a priori probabilities or odds, in advance of the occurrence of the next stimulus During the subsequent reaction time new information, from successive stimulus samples, is used to modify these initial odds until the posterior odds become sufficiently large (or small) to warrant the risks involved in making a decision, at which time a response is initiated The fact that a computer can be programmed to make decisions in this manner, and that such a computer would show differential reaction times and make proportions of errors much like human *Ss* lends some degree of plausibility to the sequential stimulus sampling theory

The general conclusions reached in regard to cognitive set learning and skilled performance are as follows (*a*) cognitive sets develop very slowly, and tend to generalize over many classes of similar situations, (*b*) however, once established they can be "called up" in a matter of about a second or less by an appropriate cue, (*c*) elicitation and utilization of cognitive sets is often facilitated by the availability of verbal labels for use as cues although such labels are not necessary

The experiments discussed above involved discrete and serial vocal and motor responses It is assumed that similar phenomena can be found in many other perceptual motor tasks, such as a batter facing a pitcher who can throw a variety of pitches, and a pilot learning to handle an airplane in rough air In both of these examples stimulus probability learning is involved

Some Cross-Category Problems of Skill Learning

In the remainder of this paper the relevance of the general views regarding skilled performance and skill learning processes discussed up to this point will be considered as they apply to several additional and more specific topics Emphasis will be on the relation of skill learning to other learning tasks

Discrimination Capacity and Skill Learning

The first of these topics is the relation of sensory, perceptual, and short-term memory processes to skill learning It is proposed that precision in controlling the amplitude and timing of a movement is limited by, and should never exceed, the capacity of an individual to learn to discriminate the external and proprioceptive feedback resulting from the single responses or series of responses Supporting this conjecture is the fact that augmentation of feedback, in such a way as to increase discriminability, usually results in reliable increments in skilled performance and learning rate This is true when the magnitude of display changes is increased so as to enhance the discriminability of visual cues, or when the elasticity or viscosity of a control is increased so as to enhance the discriminability of proprioceptive feedback Bahrick (1957) has shown, for example, that addition of a spring to a control leads to improvement of blind positioning movements of that control—presumably because each increment in response amplitude now has associated with it a more easily discriminated change in force, resulting from displacement of the spring Bahrick and Noble (1961) have recently obtained empirical gradients of response generalization, which they interpret as reflecting failure in discriminating among adjacent responses, i e , response similarity effects, and have studied the effects of increasing stimulus vs increasing response discriminability They conclude that response similarity has a greater effect upon response generalization than does stimulus similarity in a task where different motor responses are made to a series of stimuli

Where the same response, or a small set of responses, are made over and over again, it would appear, however, that the average timing and precision of responses may come to exceed the *S*'s ability to discriminate single stimuli, at least as discriminability is measured by standard psychophysical methods Such evidence suggests that the *S* is able eventually to learn somehow to adjust his responses in accordance with the accumulated information gained from a series of discriminations, and to reduce his response variability below his error in judging single responses In other words, at very high levels of practice in a highly coherent task, variability of response becomes less and response differentiation better, than would be predicted from ordinary psychophysical data taken early in learning Perhaps this simply indicates a capacity for long continued discrimination and probability learning, similar to the demonstrated capacity for long continued improvement in the speed of performance Another important fact to consider, however, is that the criterion for discrimination often is a 75% or even higher proportion of successes (in 2-choice tasks) whereas

probability learning effects may continue to operate at levels only very slightly above chance, albeit over a very large number of responses

As an aside, at this point, learning theorists who place great emphasis on the role of response produced stimuli need to be reminded that discrimination of such stimuli is relatively poor if the judgment is on an absolute basis along a single dimension For example, *Ss* identify a maximum of only about 2 bits of information in judging the pressure which they are applying to a control Again, this does not rule out the possibility of slow probability learning effects, however

The importance of stimulus discrimination learning varies, of course, with the nature of the overall task For example, proprioceptive discrimination should be most important in skills where there is little uncertainty regarding what response is appropriate and *S* is confronted with the task of precise response execution Learning to execute a complex dive would be an example

Short-Term Memory and Skill Learning

Several writers have discussed the role of short term running memory in the learning of perceptual motor skills Again we find in the early work of Cattell (1886) a suggestion of the effects of memory limitations on the speed of perceptual motor performance Cattell exposed letters which the *S* read as they passed behind a slit By varying the width of the slit *Ss* were permitted to see one or more than one letter at a time—a condition called *preview* He found that ability to see more than one letter at a time increased serial reading speed up to a preview limit of four or five letters, but that the major advantage was gained as soon as the second letter was visible Wagoner and Fitts, in an unpublished study completed several years ago, found similar results for preview when the task was to push the appropriate one of five keys at exactly the time that light points, moving downward in five different columns, passed behind a horizontal line which extended across the five columns Most of the advantage gained from permitting preview was achieved by viewing the next light in the sequence A preview of 12 oncoming lights was little better than the preview of only one

Crossman (1960) has recently published a theoretical analysis of ' perceptual anticipation in pursuit tracking as a function of a hypothetical limit on capacity for processing information, and a limit on capacity for running memory His main thesis is that one or the other of these capacities may determine the limit of performance in a given situation For example, when redundancy is high but the task is very complex there should be considerable gain from preview over several seconds or over a large

number of forthcoming stimuli, memory capacity presumably would be the limiting factor in this situation.

Continuous and serial tasks often provide an opportunity for measuring the lag, and hence the memory load, between the information-seeking responses of the eye, and the output responses of the hand or vocal mechanism. Studies of eye-hand span in typing, playing the piano, and assembly work, studies of eye voice span in oral reading, and studies of lag in copying telegraph code all indicate that closely comparable and near maximum amounts of information are carried in running memory in different perceptual-motor tasks.

Such limitations as these become especially important when we turn to the mechanisms which may account for continued improvement in highly coherent, speed skills at advanced levels of practice. These will be considered next.

Continued Improvement in Skills

Crossman (1959) is the only worker who has addressed himself specifically to the development of a theory regarding how continued improvement comes about at very high levels of skill and why improvement continues so slowly over so long a period of time.

Crossman assumes a general probability learning model. Thus he states that "the operator can be imagined to possess a repertoire or stock of *r* different methods, from which he picks one by chance for each cycle" (1959, p. 159). Each method has an initial probability, and these probabilities are modified after each (discrete) response as a function of feedback. However, since all responses at high levels of skill are "successful" in the ordinary sense that they all permit attainment of the objective, a critical problem is to account for the mechanism by means of which the time required for specific responses is discriminated with the degree of precision required by the selective process. One interesting hypothesis, among several advanced by Crossman, is a mechanism based on the rapid initial decay of short-term memory. Feedback from faster responses is hypothesized to arrive after less decay of the memory process than does feedback from slower responses and thus on the average should provide a positive increment in probability of the subsequent use of the faster response. Here again, it is important to remember that the selective process operates over very long periods of time, and hence may be able to use relatively unreliable or "noisy" feedback information.

Another source of evidence regarding long term skill learning, mentioned briefly in an earlier section, concerns the ability of *Ss* to learn to carry on more than one task simultaneously. Bahrick, Noble, and Fitts (1954) reasoned that if proprioceptive feedback does become increasingly

important after extended learning in highly coherent tasks, and if less and less reliance on visual cues is necessary as learning of a first task progresses, then a second visual task should have less and less interfering effects on the first task to the extent that the first task is (*a*) highly coherent, and (*b*) highly overpracticed Both ideas were verified by experiments These studies also provided unequivocal evidence that learning had been going on long after the original criterion employed in measuring learning on the first task had ceased to indicate evidence of improvement

Feedback Variables in Skill Learning

Bilodeau and Bilodeau, in their recent review of motor-skills learning, conclude that "studies of feedback or knowledge of results show it to be the strongest, most important variable controlling performance and learning" (1961, p 250), and provide a comprehensive review of the effects of this variable In spite of its obvious importance, I shall not have much to say about the topic of feedback, however, partly because the evidence is so clear, and partly because, from a theoretical standpoint, there appears to be little in this area that is peculiar to skill learning as distinct from other learning tasks

One interesting line of work on feedback should be mentioned, however This is the effect of augmented feedback, the availability of special feedback information which ordinarily is not present in a skill learning task Some of the effects of augmented feedback appear to be motivational Fitts and Leonard (1957), for example, found that a continuous series of clicks at the rate of 2 per sec heightened performance in a speeded perceptual task Subsequent studies (Smode, 1958, Kinkade, 1960) suggest the possibility of additional learning effects as well as performance effects

One of the avenues by means of which augmented feedback may influence skill learning is through its effect on discrimination learning, especially during advanced stages of practice When the task is a relatively incoherent one, augmented feedback may also provide much more reliable information than is ordinarily available to the *S* as regards the adequacy of his general strategy or cognitive set In other words, as learning approaches the hypothetical limit of the *S*'s ability to discriminate available feedback, in either coherent or incoherent tasks, augmented feedback may become increasingly useful because of the more precise information it provides to the learner

Skill Learning and Problem Solving

Several writers have recently emphasized the analogy between thinking and skill learning At first these two forms of behavior may seem to be at

opposite ends of a continuum However, the approach to skilled performance developed in the present paper, especially the emphasis on the patterning and organization of skills and on the importance of cognitive set phenomena, make the relationship more plausible Bartlett (1958) stressed the idea that the extrapolation and orderly sequencing of responses is involved in both skilled performance and in thinking In a similar vein Piaget (1950) proposed that "intelligence sensori motoric" precedes and provides the foundation for "intelligence intellectualle"

Certainly the role of cognitive set in skilled performance, and the general adaptive system model of skilled performance, with its emphasis on hierarchical programs, brings skilled and problem solving behavior closer together than would have been commonly proposed a few years ago Perhaps there the matter should rest until additional evidence regarding the relationship is at hand

Some Relatively Unique Aspects of Skill Learning

Lest the similarity of skill learning to other learning tasks be overemphasized, I shall close this paper with brief mention of four areas that are relatively unique to, but very important for, the study of skilled processes These topics will not be discussed at length because they are judged to be of somewhat peripheral interest to those working in other areas of human learning

Open-loop behavior at various stages of learning —For various theoretical reasons it would be especially instructive to study perceptual motor performance in continuous or serial tasks where either the feedback which *S* normally expects and uses or the input which is mixed with feedback, is temporarily eliminated so that *S* would respond to one of these sources of information alone For example this would permit a determination of the relative importance of input vs feedback, and of the degree to which response-generated information (feedback) is sufficient to maintain performance In the area of skills this is called the study of "open loop" responses It is an easy matter to eliminate visual or auditory feedback in most perceptual-motor tasks (such as by having subjects close their eyes), but the elimination of proprioceptive feedback is usually not possible, and as long as the latter is present *S* quickly realizes that he is no longer getting exteroceptive feedback and usually changes his behavior accordingly Some success has been met in eliminating the input and arranging the task so that *S* responds only to his own feedback However it is difficult to do this without the *S* realizing that the input has suddenly stopped Thus the separate effects of input vs feedback are only partly understood

Analysis of complex forms of stimulus response congruence —Some of

the highest levels of skill are attained in tasks in which one movement pattern is superimposed upon another pattern, such as in aiming a gun at a moving target while at the same time maintaining an upright posture while standing on a moving platform, or in throwing a football at a moving target while running at full speed Performance in such tasks would appear to exceed ordinary human information processing capacity and must therefore depend on the use of highly overlearned or automatized subroutines It would be highly instructive to study such learning but little work has been done on this problem, perhaps because of the technical difficulties involved in recording and analyzing separately two concurrent response processes

Speed-accuracy tradeoff—Man has the rather unique ability to exchange speed for accuracy of responses and vice versa High levels of ability in effecting such compromises are evidenced in almost all information handling and control tasks, suggesting a basic interdependence between mechanisms for the regulation of timing and mechanisms for the regulation of the direction and amplitude of movement The sequential stimulus sampling theory described previously is one example of an effort to understand these interrelations A recent study by Fitts and Peterson (1964) provides clear evidence that the speed accuracy relation is determined by motor centers which are quite separate from those involved in the control of choice reaction time, since movement time and reaction time were found to be quite independent functions

Human transfer functions—A small but active group of researchers have for the past ten years been interested in analyzing and constructing mathematical models of the processes by means of which a man learns to control the output of a complex dynamic system, such as one in which the output is the second or third integral of the input This is a problem unique to the study of skill learning (Licklider, 1960) The topic is a highly interesting one because, as mentioned earlier, the same theory which has been developed for the analysis and synthesis of dynamic physical systems can, to a considerable extent, be applied directly to the description of human perceptual motor learning and performance Theory of sampled-data systems is especially relevant to current work in this area

Summary

In this chapter I have placed particular emphasis upon (*a*) the importance of research on spatial temporal patterns of behavior, including patterns lasting for only a second or less, and (*b*) the importance of stimulus coherence as the objective characteristic of tasks and sequences which appears to be of widest importance and of most theoretical interest

The theoretical idea which I would re-emphasize is that of hierarchical

processes One useful way to view the higher level processes in a complex skill is by analogy with the executive routines written into computer programs The corresponding view of lower level processes is by analogy to the loops and subroutines of such programs The theoretical view of skilled performance here proposed minimizes the role of motor behavior per se, and thus removes the principal basis for the commonly made distinction between verbal and motor processes Instead it places major emphasis on the intrinsic coherence of stimulus and response sequences and the cognitive or higher-level processes that govern behavior sequence, and suggests that important aspects of response sequences include such factors as timing, the interrelations of speed, accuracy, and uncertainty, and the limitations imposed by capacities for discrimination and memory

The crucial point in developing a general theory of skilled performance, and in support of the view that verbal and motor processes are highly similar, is the conclusion that skilled performance is dependent on discrete or quantized processes Thus the study of discrete perceptual motor responses, including the study of reaction time, movement time, and response accuracy (errors), can be viewed as contributing to an understanding of serial and continuous communication and control skills on the one hand and to an understanding of the organization of thinking decision making, and verbal behavior on the other hand

References

ALLUISI E A Conditions affecting the amount of information in absolute judgments *Psychol Rev* 1957 64 97–103

BAHRICK, H P An analysis of stimulus variables influencing the proprioceptive control of movement *Psychol Rev* 1957 64 324–328

BAHRICK H P NOBLE M E & FITTS P M Extra task performance as a measure of learning a primary task *J exp Psychol* 1954 48, 292–302

BAHRICK H P & NOBLE M On stimulus and response discriminability *J exp Psychol* 1961 61, 449–454

BARTLETT F C *Thinking* London Allen & Unwin 1958

BILODEAU E A & BILODEAU I McD Motor skills learning *Annu Rev Psychol* 1961 12, 243–280

BLACKBURN J M Acquisition of skill an analysis of learning curves Great Britain I H R B Report No 73 1936

BRAINARD R W IRBY T S FITTS P M & ALLUISI E A Some variables influencing the rate of gain of information *J exp Psychol* 1962 63 105–110

BRYAN W L & HARTER N Studies on the telegraphic language The acquisition of a hierarchy of habits *Psychol Rev* 1899 6, 345–375

CATTELL J McK The inertia of the eye and brain *Brain* 1885 8, 295–312

CATTELL, J McK. The time it takes to see and name objects *Mind* 1886 11, 63–65

CRAIK K J W Theory of the human operator in control systems II Man as an element in a control system *Brit J Psychol* 1948 38, 142–148

CROSSMAN, E R F W A theory of the acquisition of speed skill *Ergonomics*, 1959, 2, 153–166

CROSSMAN, E R F W The information capacity of the human motor system in pursuit tracking *Quart J exp Psychol*, 1960, 12, 1–16

FITTS, P M The information capacity of the human motor system in controlling the amplitude of movement. *J exp Psychol*, 1954, 47, 381–391

FITTS P M Skill training In R. Glaser (Ed), *Training research and education* Pittsburgh Univer Pittsburgh Press 1962

FITTS P M, BAHRICK, H P, NOBLE, M E, & BRIGGS, G E *Skilled performance* New York Wiley in press

FITTS P M, BENNETT, W E., & BAHRICK, H P Application of autocorrelation and cross-correlation analysis to the study of tracking behavior In G Finch and F Cameron (Ed), *Air Force human engineering personnel and training research* Washington National Academy of Sciences—National Research Council Publ 455, 1956 pp 125–141

FITTS P M., & DENINGER R. L S–R compatibility Correspondence among paired elements within stimulus and response codes *J exp Psychol*, 1954, 48, 483–492

FITTS P M, & LEONARD J A Stimulus correlates of visual pattern recognition—a probability approach. Final report Contract No NONR-495(02) Columbus Ohio Ohio State University Research Foundation Oct., 1957

FITTS, P M., & PETERSON J R The information capacity of discrete motor responses *J exp Psychol*, 1964, in press

FITTS P M., PETERSON, J R. & WOLPE, G Cognitive aspects of information processing II Adjustments to stimulus redundancy *J exp Psychol*, 1963, 65, 423–432

FITTS P M, & SEEGER, C M S–R compatibility Spatial characteristics of stimulus and response codes *J exp Psychol*, 1953, 46, 199–210

FITTS P M, & SWITZER, GAIL Cognitive aspects of information processing I. The familiarity of S–R sets and subsets *J exp Psychol* 1962 63, 321–329

FLEISHMAN E. A., & HEMPEL, W E Changes in factor structure of a complex psychomotor task as a function of practice *Psychometrika* 1954, 19, 239–252

FLEISHMAN, E A, & HEMPEL, W E. Relation between abilities and improvement with practice in a visual discrimination reaction task. *J exp Psychol*, 1955 49, 301–312

FULLERTON G S., & CATTELL, J McK. *On the perception of small differences, with special reference to the extent, force, and time of movement* Philadelphia Univer Pennsylvania Press, Philosophical Series No 2, 1892

GIBSON J E Making sense out of the adaptive principle *Control Engng*, 1960, 7, 113–119

GREW S Information gain in tasks involving different stimulus-response relationships *Nature*, 1958 182, 1819

HICK, W E. Discontinuous functioning of the human operator in pursuit tasks. *Quart J exp Psychol* 1948 1, 36-57

HICK, W E. On the rate of gain of information. Quart. *J exp Psychol.*, 1952, 4, 11–26

HIRSH I J & SHERRICK C E. Perceived order in different sense modalities *J exp Psychol*, 1961 62, 423–432.

HYMAN R. Stimulus information as a determinant of reaction time. *J exp Psychol.*, 1953 45, 188–196

KELLER, F S The phantom plateau *J exp Anal Behav*, 1958, 1, 1–13

KINKADE, R G The effect of augmented feedback on the acquisition and transfer of a tracking skill Unpublished Ph D thesis, Ohio State Univer, 1960
KLEMMER, E T Communication and human performance *Human Factors J*, 1962, 4, 75–79
LASHLEY, K S The problem of serial order in behavior In L A Jeffress (Ed), *Cerebral mechanisms in behavior* New York Wiley, 1951
LICKLIDER, J C R Quasi linear operator models in the study of manual tracking In R D Luce (Ed), *Developments in mathematical psychology* Glencoe, Illinois Free Press, 1960
MILLER, G A The magical number seven plus or minus two Some limits on our capacity for processing information *Psychol Rev*, 1956, 63, 81–97
MILLER, G A, GALANTER, E, & PRIBRAM, K H *Plans and the structure of behavior* New York Holt, 1960
MONTPELLIER, G de Note sur l'acceleration dans les mouvements volontaires de la main *Arch Psychol, Geneve*, 1937, 26, 181–197
NEWELL, A, SHAW, J C, & SIMON, H A Elements of a theory of human problem solving *Psychol Rev*, 1958, 65, 151–166
NEWMAN, E B Men and information A psychologist's view N 2 del Supplemento del Neuovo Cimento, 1959, 13, 539–559
NOBLE, M E, FITTS, P M, & WARREN, C E The frequency response of skilled subjects in a pursuit tracking task *J exp Psychol*, 1955, 49, 249–256
PIAGET, J *Psychologie de l intelligence* New York Harcourt, Brace, 1950
PIERCE, J R, & KARLIN, J E Reading rates and the information rate of a human channel *Bell Syst tech J*, 1957, 36, 497–516
RIESEN, A H, & AARONS LOUISE Visual movement and intensity discrimination in cats after early deprivation of pattern vision *J comp physiol Psychol*, 1959, 52, 142–149
RUHLE, W, & SMITH, K U Effects of inversion of the visual field on human motions *J exp Psychol*, 1959, 57, 338–343
SMOOE, A Learning and performance in a tracking task under two levels of achievement information feedback *J exp Psychol*, 1958, 56, 297–304
SNODDY, G S Learning and stability *J appl Psychol*, 1926, 10, 1–36
SNYDER, F W, & PRONKO, H H *Vision with spatial inversion* Wichita, Kansas Univer of Wichita Press, 1952
STETSON, R H A motor theory of rhythm and discrete succession *Psychol Rev*, 1905, 12, 250–370
TAYLOR, F V Psychology and the design of machines *Amer Psychologist*, 1957, 12, 249–258
TROLAND, L T *The fundamentals of human motivation* New York Van Nostrand, 1928
UNDERWOOD, B J, & SCHULZ, R W *Meaningfulness and verbal learning* New York Lippincott, 1960
WELFORD, A T The psychological refractory period and the timing of high speed performance—A review and a theory *Brit J Psychol* 1952, 43, 2–19
WOODWORTH, R S The accuracy of voluntary movement. *Psychol Rev Monogr Suppl*, 1899, Whole No 13
WOODWORTH, R S *Dynamics of behavior* New York Holt, 1958

The Generality of Research on Transfer Functions

COMMENTS ON PROFESSOR FITTS' PAPER

G E BRIGGS
Ohio State University

Dr Fitts has provided an excellent coverage of what can be called the discrete case in perceptual motor skill learning In this short discussion some information on the continuous case will be presented Specifically, research and theory will be described which are based on derivations of transfer functions for the human operator in continuous control or tracking tasks

The area of servo engineering received considerable support starting with World War II, and the result has been a rapidly developing sophistication in the analysis and design of automatic control devices for both continuous and discrete systems One of the basic analytical tools developed in this area is the transfer function methodology Formally, a transfer function is defined as the ratio of the output from a "black box" to its input. More useful to us, the transfer function is an analytic statement which predicts an output from operations imposed on the input In most of the work, a linear differential equation has served as the basic analytic statement

HISTORICAL PERSPECTIVE

It is only natural that this methodology would be applied to quantify performance by the human black box when he controls a vehicular system All that is required is a model of the human operator which can be fitted to the same input signals to produce outputs as similar as possible to those from the human tracker The notion that a linear equation could be fitted to human tracking behavior was first proposed by Tustin (1947) and it grew out of the very active war time research in England on servomechanisms for the control of large gunnery equipment Essentially, the equation or model used was a first-order differential with a transmission lag

$$p\theta_L = (k_1\epsilon + k_2 p\epsilon)\, e^{-p\tau} \qquad (1)$$

which represented a relationship between the rate characteristic of the human operator's output ($p\theta_h$) and his input (ϵ or tracking error) There are three constants in this equation T which is analogous to human reaction time, and K_1 and K_2 which are of primary importance since they can be interpreted as indicants of how the operator weighted tracking error amplitude (K_1) and tracking error rate (K_2) in generating his output

In other words, when the equation is fitted to tracking data, K_1, K_2, and T are adjusted so that a minimal residual variance is left unexplained A great deal of effort has been expended by researchers in this field in an attempt to reduce the residual variance or remnant, as it is called One of the techniques used has been to employ more terms on the left of the above equation

$$K_a\theta_h + K_b p\theta_h + K_c p^2\theta_h = (K_1\,\epsilon + K_2 p\,\epsilon)\, e^{-pT} \qquad (2)$$

Essentially this is the function utilized by Ornstein (1961) and it is comparable to that equation employed by McRuer and Krendel (1957) in their monumental review and reanalysis of work in this area A more elaborate version was employed by Fuchs (1962) which is the same as Eq 2 with the exception that a second order differential equation appears on the right

$$K_a\theta_h + K_b p\theta_h + K_c p^2\theta_h = (K_1\,\epsilon + K_2 p\,\epsilon + K_3 p^2\,\epsilon)\, e^{-pT} \qquad (3)$$

Perhaps the most elaborate attempt in this area was undertaken by Cacioppo, Mayne, Mead, and others at the Goodyear Aircraft Laboratories (see 1955) These researchers utilized essentially the same transfer function as that listed in Eq 3 but added several nonlinear operations to the model including a 'threshold," 'dither," and "anticipation " Apparently, the anticipation operation did more to account for the remnant than did either of the other two nonlinear operations Essentially, this operation was inserted to simulate the tendency of the human operator to reverse the direction of his control movement when the tracking error signal changes sign (to respond in an "alert manner to future movement requirements)

Throughout all of this has run the assumption that the human tracker could be described by a quasi linear transfer function McRuer and Krendel (1957) indicate that for simple tracking task dynamics, such quasi linear functions explain 34 to 99% of the variance, while more complex dynamics resulted in fits which accounted for between 25 and 88% of the variance Performance in the simple tracking tasks was particularly well accounted for since while the lowest figure cited above is 34%, the majority of such research has explained 80 to 99% of the variance It follows from this that the methodology is quite powerful How are these analyses implemented?

The Techniques

There have been several techniques used to fit transfer function equations to tracking data The original technique employed Fourier analyses to match input and output frequencies and examine phase differences, subsequently, auto- and cross-correlation analyses were applied with good results These techniques are described by McRuer and Krendel (1957) A variation of the correlation technique was developed by Knowles, Holland, and Newlin (1957) who employed a multiple regression analysis to predict human output (control stick acceleration) from tracking error amplitude, rate, and acceleration and from stick position, velocity and acceleration This latter technique is more familiar to psychologists, however, there was a considerable amount of variance left unaccounted for from 43 to 48%

Perhaps the most intriguing technique was developed by Cacioppo et al at the Goodyear Aircraft Corporation and was employed in a series of studies (Goodyear Aircraft Corp , 1950, 1952, 1953, 1955) This represented a radical but logical departure from the original approaches, and it is a technique which provides remarkable efficiency

Basically, this technique involves wiring an analog computer to perform the derivative operations listed in Eq 2 or 3 plus the insertion of the reaction time lag dither, threshold and anticipation functions on the computer In fact, the technique utilizes an analog of the human operator —a set of information processing functions and a set of nonlinear operations which are assumed to be analogs of such functions and operations within the human "black box"

The human and the analog are exercised simultaneously, the human tracker observes a visual display of tracking error and tries to null that signal (ϵ) by generating forces against the control stick (θ_k), while the analog tracker has the same error signal presented in the form of a time-varying voltage and attempts to reduce that voltage to zero by providing its output signal The two output signals are summed and the experimenter adjusts the various constants in the analog to provide a minimum difference between human output and analog output The Goodyear technique was rather successful in matching analog to human operator, and one of the more interesting criteria used to determine the adequacy of the fit was to switch off the man's output and substitute the analog output without informing the human tracker when this had occurred The authors report that the human would not be aware that he has been replaced by the analog for periods of time up to 20 or 30 sec!

A major shortcoming of this "manalog" technique was that the experi

menter had to adjust the constants manually by a heuristic process Fuchs (1962) used a semiautomatic variation of this technique which fitted one constant at a time, and Ornstein (1961) developed the most sophisticated version of the manalog technique which fits all constants in the transfer function simultaneously by the principle of steepest ascent

One conclusion that could be reached is to say So what? The manalog technique produces an automatic device which when adjusted properly will provide controlling signals (an output) which resemble the controlling signals of a human tracker Is this anything more than a demonstration that talented engineers and psychologists can simulate human information processing in a continuous control task? I believe that work is of considerably more import to psychology than would be a mere demonstration The reason I believe this is that the constants of the manalog and the constants of the transfer functions utilized by the others working in this area behave systematically as a function of changes made in the task environment and in the human operator himself

Some Empirical Data

Russell (1951) provided his *Ss* with either low, medium, or high speed inputs The behavior of the transfer function constants, particularly K_e was quite systematic the higher the input speed, the lower the weight associated with error rate This is a logical finding in that with high speed inputs a tracker frequently 'falls behind the input—he quickly gets out of phase," and the only way to get back in phase is to adjust for an observed amplitude of error With more slowly changing inputs the tracker can stay "on target' more consistently, and he must do so by correctly matching input rates of change with his output θ_k—he must utilize error rate information

Walston and Warren (1954) provided either a conpensatory or a pursuit tracking task and found that the K_2 constant was greater for the pursuit task than for the compensatory task Again this is logical in that a pursuit display more obviously provides rate information than does a compensatory display In one of the Goodyear studies (1955) the transfer function constants for the rate and acceleration aspects were considerably higher for a more experienced pilot than for a less experienced pilot In fact, the K_2 constant for the more experienced man was twice that for the pilot with less training and experience, while the acceleration constant K_3 of Eq 3, was 33% greater

These observations led Fitts et al (1959) to propose the hypothesis that as a man gains skill in a particular tracking task one would find increasing weight on the rate and/or acceleration components of his transfer

function with diminishing weight being associated with the error amplitude component This hypothesis was tested and confirmed by Fuchs (1962)

Time does not permit a more extensive recital of empirical data I only wish to point out that changes in the input characteristics, changes in the display characteristics, and changes within the human operator bimself are all reflected by systematic and predictable changes in the constants of a human transfer function I might add that Ornstein (1961) showed a similar systematic change in K_2 when he manipulated the characteristics of the control device and the kinds of feedback signals provided to S

In other words, one has a technique here which reflects human behavior in a most analytic way and does so quite systematically The approach is so consistent with logic and with independently derived conclusions about behavior in continuous control tasks that it not only proposes reasonable kinds of information processing operations utilized by the human operator but also provides quaotitative indices of how the operator utilizes these functions or something analogous in generating his output It gives one the peculiar feeling that he has managed to get inside the S without the use of a scalpel or a set of micro-electrodes

A Criticism

It hardly seems logical that one who is convioced of the importance and power of the transfer function methodology should want to criticize the work at the same time he is suggesting it as a means of quantifyiog behavior in other areas of learning research I will do so, bowever, because if my suggestion is adopted (and it appears to be eveo before I so proposed it), I fear that the same error may be committed agaio

Actually, my criticism is in regard to an oversight more than to an error Specifically, a great deal of time and energy has been spent since Tustin in developing ever more complex and efficient ways to fit transfer functions Parallel to this has been a dominant motive to reduce the nonlinear remnant to smaller and smaller values I applaud these motives and efforts up to a point, that point being where one is more interested in the technique than in the use to which it can be placed the descriptive quantification of behavior and the analytic tests of hypotheses about behavior We have had far too little of the latter and therefore I criticize the work on this ground

Of course one wants as reliable, valid, efficient, and powerful a research methodology as he can get, but in the final analysis I think the psychology of skill learning would be much further advanced had an extensive program of research been carried out five or six years ago, with the admittedly less-than perfect techniques available, which was devoted to the *use* rather than the further development of transfer function methodology

This did not happen then and it has not happened since, thus, one finds some rather large gaps in empirical data which are quite provoking when one attempts to develop a theoretical understanding of continuous skill processes

Relation to Other Human Learning

What does the transfer function methodology have to offer us here in our attempt to relate and integrate the several areas of human learning? I have suggested that in skill research this methodology provides for a useful description and quantification of information processing functions I think the same kind of description and quantification can be obtained if the methodology is applied to information processing functions in the problem solving, decision making and concept formation areas In fact, there have been encouraging initial attempts to use analogous techniques in the discrete case Kendler, in this volume (pp 223–225), touched upon computer simulation of cognitive processes and Newell, Shaw, and Simons (1958) report on the General Problem Solver program for digital computer simulation of human problem solving behavior

I foresee some problems for the discrete case For the analyses of tracking data the researchers had a basic model, the general servo paradigm, and a rich library of concepts I do not know of a comparable general model for behavior in the discrete case, and this lack can hold back any area of scientific inquiry Certainly there is no lack of concepts, as indicated by Fitts' lucid exposition here, but where is the unifying model?

Summary

A brief description has been given of behavioral analysis by the transfer function methodology as applied to the continuous (tracking) case Covered were the general quasi linear techniques and some of the empirical data An attempt was made to show that this methodology (*a*) is quite analytical and powerful, (*b*) gives behavioral data which are logically related to a variety of independent variables, and (*c*) offers guidance to those who wish to perform comparable analyses of behavior in the discrete case

References

Fitts P M Bahrick H P Noble, M F & Briggs G E Skilled performance Chap 12 Ohio State University Res Found Final Report Contract AF 41 (657) 70 March 1959

Fuchs A H The progression regress on hypotheses in perceptual motor skill learning *J exp Psychol* 1962 63, 177–182

Goodyear Aircraft Corp A proposal to study the dynamic characteristics of man

as related to the man aircraft system *Goodyear Rep*, April 1955, No GER-3006 A

GOODYEAR AIRCRAFT CORP Final report human dynamic study *Goodyear Rep* April 1952, No GER-4750

GOODYEAR AIRCRAFT CORP Investigation of vestibular and body reactions to the dynamic response of a human operator final report *Goodyear Rep*, November 1953, No GER 5452 (BuAer Rep AE 61 6)

GOODYEAR AIRCRAFT CORP Investigation of control "feel' effects on the dynamics of a piloted aircraft system Goodyear Rep, April 1955, No GER 6726

KNOWLES, W B, HOLLAND, J G, & NEWLIN, E P A correlational analysis of tracking behavior *Psychometrika*, 1957, 22, 275–287

MCRUER, D T, & KRENDEL, E S Dynamic response of human operators *USAF WADC tech Rep*, October 1957, No 56 524 (ASTIA No AD 110693)

NEWELL, A, SHAW, J C, & SIMON, H A Elements of a theory of human problem solving *Psychol Rev*, 1958, 65, 151–166

ORNSTEIN, G N Applications of a technique for the automatic analog determination of human response equation parameters *North American Aviation Rep* January 1961, No NA61H-1

RUSSELL, L Characteristics of the human as a linear servo element Unpublished master's thesis, Massachusetts Institute of Technology, 1951

TUSTIN, A The nature of the operator s response in manual control and its implications for controller design *J Inst electr Eng*, 1947, 94, Pt IIA, 190–207

WALSTON, C E, & WARREN, C E A mathematical analysis of the human operator in a closed loop control system *USAF Personn Train Res Cent res Bull*, December 1954 No 54–96

Problem Solving

ROBERT M GAGNE[1]
Princeton University

The field of experimental investigation of problem solving presents an aspect of great diversity in the materials and techniques employed for its study, yet at the same time a relatively circumscribed number of memorable ideas The phrase "problem solving" usually makes one think, initially, of the work of Ruger (1910) on the solving of puzzles, of Maier (1930, 1933) with various practical situations like the hatrack and the pendulum problems, of Duncker (1945) with paper clip and pyramid problems of a somewhat similar nature, of Luchins' (1946) water jar problems, of Rees and Israel's (1935) anagrams, of Katona's (1940) card trick and matchstick problems and of Wertheimer's (1945) parallelogram Coming down into more modern times, there have been adaptations of anagram and other kinds of word problems, the use of sequential problems like Twenty Questions, troubleshooting and the Psi Apparatus, a variety of numerical problems, and the solution of problems like "oddity" and "non oddity" by monkeys, as well as reversal shift problems by children I have not tried to mention all of the kinds of problems which have been studied, but simply those which seem to have prominence when one surveys the field from memory From this point of view, it is a field at once highly differentiated and rather sparsely occupied

What does learning have to do with problem solving? A first answer to this question might well be that, for one thing, the solving of a problem is a set of events which must have been *preceded* by learning In order to solve anagrams, one must previously have learned words and their meanings, to solve water-jar problems, one must have learned to add and subtract numbers, to solve a pendulum problem, one must have learned about pendulums, and perhaps other things as well

This point of view toward the relation of learning to problem solving appears to be perfectly valid But it is probably not the one most profitable to examine here Instead, I should like to consider the proposition that problem solving, regardless of what may have preceded it is itself a form of learning All of the necessary formal properties of learning

[1] PRESENT ADDRESS American Institute for Research Pittsburgh, Pennsylvania

would seem to be applicable to the problem solving situation An individual organism is placed in a situation over which an experimenter has control Stimuli are presented, perhaps in some particular order, but at any rate extending over a period of time The individual responds, and his performance after the experimental treatment is found to be altered from what it was before, there is a measurable change from an earlier point of time in the experimental session to a later point of time Furthermore, it is a change which can be shown to be retained Now, if he chooses, the experimenter may wish to make the inference that the subject (*S*) "has performed a successful act of thinking" It seems equally reasonable, however, to make the inference that the observed change means that the *S has learned* If an experimenter makes these kinds of observations on the *S*, he may prefer to call it an experiment on problem solution, at the same time, his hesitation to employ the same *S* over again with the same categories of problems provides ample proof that he thinks the *S* has learned something

If one accepts the observed change in performance in a problem solving situation as evidence that learning has occurred, then it becomes of interest to inquire "what kind of learning?" Is this a matter of operant conditioning, perhaps of chains of responses? Is it discrimination learning? Is it a particular variety of paired associate or serial learning? Or, if it is not any of these, how is it related to these other forms of learning? With respect to this last question, it may be noted that Underwood (1952) has discussed several methods of determining the relationship between problem solving and other forms of learning In general, the present discussion may be said to be most compatible with the third of his methods, in its attempt to identify the manipulable variables in problem solving, and perhaps also to suggest what might be the functional relationships between these variables and the behavior observed

My purpose will be, then, to discuss the variables which appear to govern problem solving behavior as a form of learning, with the aim of defining them This will be done by considering first the dependent variables and then the independent variables The attempt will be made also to point out the identities, similarities, or differences in these variables as they occur in the problem solving situation, compared with other learning situations In addition some speculations concerning intervening processes will be offered on the basis of the analysis just outlined

Dependent Variables in Problem Solving

The measurement of problem solving performance is not an easy matter, and considerable attention has been devoted to it by various investigators

Ray (1955) summarizes a variety of attempts which have been made to quantify behavioral measures in this area. Duncan (1959) points out that sometimes measures appear to have the purpose of identifying process rather than product. Quite often, too, there appears an emphasis on the *quality* of the performance, as in Wertheimer's (1945) treatment of 'good and "stupid' solutions, Duncker's (1945) references to "meaningful and "nonsensical" errors, and Maier's (1930) descriptions of different categories of solution. The replacement of such qualitative scores, often dependent for their usefulness on reliability of judgments among scorers, with objectively determinable measures would seem to be an eminently desirable aim for studies of problem solving.

As is the case with other types of learning, the investigation of problem solving must be concerned with two general categories of behavioral measures: (1) rate of attainment of some criterion performance, and (2) degree of correctness of this performance. In other words, one is sometimes interested in measuring the speed with which the inferred process of learning has occurred, and in other experiments, on the goodness of the performance which has resulted from this process. In the study of problem solving, the search for fully satisfactory measures relevant to both of these categories is still going on.

Measures of Rate of Attainment

Rate of attainment of a terminal performance may be measured by such things as amount of time, number of presentations (of the problem situation or parts of it), or number of errors required for the *S* to reach some criterion. The rationale is different for each of these measures, and a careful experimenter is likely to have some misgivings about any or all of them in a given experimental situation.

Amount of time for solution.—At first glance, the passage of time would appear to be a fairly direct indicator of the rate at which learning is occurring. But in problem solving, time often appears to be a crude measure indeed. This is particularly true when there is 'unfilled' time in which no overt responses are observed, and during which the *S* may be doing nothing of consequence, or rehearsing something he previously learned, or even thinking. For this reason, it is apparent that time measures of rate of attainment often are subject to those factors which tend to produce unreliability. However, if the experimental situation is such that frequent responses are required, as when a problem demands a number of moves (Ewert & Lambert, 1932), or when a learning program is employed which requires the *S* to respond to a series of frames, it may be that elapsed time can become a satisfactory measure of attainment rate.

Number of presentations.—When one considers the use of number of

presentations of the problem or its parts as a measure, a different kind of difficulty presents itself In this case the experimenter finds it hard to decide exactly what to present on repeated occasions Does repeated presentation mean repeating instructions, repeating the exposure of stimulus materials, repeating "parts" of the problem (cf Maier, 1930)? Besides, it frequently happens that nothing gets repeated, as in studies in which only one problem presentation is made, after which the *S* is expected to come up with a solution The whole idea of using number of presentations as a measure of rate of attainment in problem solving seems somehow to be incongruous with the aims of a problem solving experiment Repetition may in fact be involved in problem solving in ways which we do not yet know about But until the question of *what* is to be presented repeatedly has received some reasonable answer, the use of this dependent variable can be seen to be not entirely free of difficulties of interpretation

A particular instance of number of presentations is the use of number of 'helps" or 'hints" during the time of problem solving, as was done, for example, in studies by Maier (1930), Reid (1951), and Gagne and Brown (1961) Although such hints may be intuitively satisfying to the experimenter, and may in fact have some effectiveness in bringing about achievement of solution, they nevertheless do not avoid the difficulty mentioned previously They constitute some "part" of the problem situation, but what part is not clear Under these circumstances, one cannot conclude that giving them the common name "hint," and then proceeding to count them, is a procedure which will result in satisfactory measurement of the rate of attainment of a problem solution

In the special case in which the learning situation is constituted of repeated presentations of the *problem as a whole,* with minimal or no instructions, the measure of number of trials to criterion has sometimes been employed as a measure of rate of attainment Such a measure is used, for example, in experiments involving transposition behavior (Kuenne, 1946) and reversal shift behavior in children (Kendler & Kendler, 1959, 1961), as well as in those concerned with oddity problems in monkeys (Harlow, 1949) In such instances, problem solving [or the "attainment of concepts" (cf Kendler, Ch 5)] is being contrasted with discrimination learning in situations in which a standard of measurement is provided by the discrimination learning When it is shown that significantly fewer trials are required to solve the problem by one group than by another (e g, by 7-year-old children vs 4-year-old children), the experimenter usually reasons that some different underlying process is involved in the two cases ("Suddenness" of solution has a long history as a criterion of problem solving) It seems clear that the measure "trials to criterion" is useful in these experiments primarily because it can be con

sidered a good indicator of rate of attainment of a discrimination The experimenter may be interested in such a measure for problem solving, particularly if he has a hypothesis about what repeated presentations are accomplishing in terms of internal processes However, it needs to be remembered that repeated presentation is only one particular set of learning conditions for problem solving, and often quite an inadequate one For the variety of problem solving situations which do not repeatedly present the stimulus situation, the use of the measure "trials to criterion" contains the difficulties previously described

Number and kinds of errors—The number of errors as well as number of errors of various types, may constitute useful measures of rate of attainment in some situations, while not in others Like 'trials to criterion,' this measure has sometimes been employed in studies designed to reveal a contrast between problem solving and discrimination An example is Kuenne's (1946) study of transposition behavior When the problem is one requiring a sequence of steps, as in the case with troubleshooting or with the Psi Apparatus (John, 1957), number of incorrect moves, number of irrational moves, or number of redundant moves, are all possible error measures The findings in one study of troubleshooting (Highland, Newman, & Waller, 1956), however, serve to emphasize the difficulties of interpretation of error scores In this study, number of moves ("checks") was found to bear little or no relation to success in solution of the problem, but there was a relation between the *pattern* of moves and successful solution Subjects successful in solving the problem were found to make more moves away from the vicinity of the malfunction site early in the session, and more moves in the vicinity of the malfunction late in the session, as compared with unsuccessful solvers Thus there may be several different reasons why excess moves are made (in other words, there are several different types of errors), and a simple relationship with rate of attainment may be too much to expect in this kind of a problem

However, there are some situations in which the degree of control exercised over the *S*'s behavior is such that the use of a score like 'excess moves' seems entirely adequate In this respect, moves may be similar to a time measure, when employed under suitably controlled conditions For example, 'excess moves' appears to be a reasonable measure of rate of attainment of the disc transfer problem employed in studies by Peterson and Lanier (1929), Ewert and Lambert (1932), and Gagné and Smith (1962) In this kind of problem, the minimum number of moves required for solution is known exactly Furthermore, the moves themselves are all of one class that is, they all involve the moving of a counter to a circle Under these circumstances, each move in excess of the minimum required may be counted as an error, and furthermore the

number of these errors may reasonably be considered to reflect a rate of attainment which permits comparability among *S*s A rationale of somewhat similar sort may be made for the use of "moves" in the Psi Apparatus, where the moves are again all of the same sort, pressing buttons

Summarizing briefly the state of affairs for the measurement of rate of attainment, we have seen that number of presentations is of limited usefulness in studies of problem solving, except in the special case where the problem situation as a whole is repeatedly presented, as when discrimination is being observed as an alternative learning process In other learning situations for problem solving, number of presentations raises the difficult problem of what is to be presented, whether parts of the problem, instructions, or hints Time and number of errors, on the other hand, have some potential usefulness as measures, but only in situations where the behavior of the *S* is suitably constrained Time, for example, appears to be a reasonable measure when frequent overt responses are demanded (as in the use of a "learning program" to bring about problem solving), thus reducing the opportunities for other activities, such as internal rehearsal, to exert their effects Errors, or number of excess moves, is similarly an appropriate measure for rate of attainment, but only when the type of error or move is comparable for different *S*s, as in problems which require frequent button pushing as part of the solution itself

Measures of Criterion Performance

In considering ways of measuring the solution of problems, it is possible to distinguish those investigations whose focus of interest is on the process of attainment or on the product, both of which types have been reviewed and discussed by Duncan (1959) Interest in process reveals itself in attempts to measure such variables as the sequence of intervening responses employed in the final performance (Ling, 1946, John, 1957), or in inferences concerning the use of strategies (Bruner, Goodnow, & Austin, 1956) The treatment here presented, however, will consider only product measures

A discussion of measurement of problem solving performance can deal fairly briefly with the matter of how single problem solution is measured But it is necessary to consider also an aspect of performance not always explicitly included in experiments on problem solving—the generalizability (or transferability) of performance Both of these kinds of measurement may be involved in the assessment of the criterion performance

Single problem solution.—The solution of a single problem is usually measured as an all-or-none score, the *S* either solves the problem or he does not. There do not appear to have been many attempts to depart from

this principle of scoring Being "partially correct" in solving a problem is considered as unlikely a state as being "slightly pregnant" The exception to this again occurs with those problems in which a number of moves of the same type is the essence of problem solution, like the Psi Apparatus and the disc transfer problem But most problems require scoring as pass or fail

A methodological implication of this scoring requirement may as well be mentioned here If pass or fail is the method of scoring one must expect to use a kind of statistical analysis which is appropriate to this measure, namely, a method which permits the testing of differences among frequencies or proportions Several possibilities, including chi square, are available to the experimenter

Generalizability of solution—It is usual, in studies of problem solving, for the pass-fail criterion to be applied to any solution which has an equivalent form For example, if the *S* has solved a matchstick problem correctly, it is not necessary that the matches assume an exact physical pattern, but only that the required number of squares be formed Or, if the *S* is solving a numerical problem, an answer of 12/5 may be considered equally correct to one of 2 2/5 This means that the "correct answer" is an abstraction from the products of responding, not the responses themselves There is a contrast here with the situation obtaining in studies of verbal learning, in which the experimenter may require the response "happy" or something with close physical similarity to it for correctness, but would not accept the response "joyful" In problem solving, a solution is considered correct when it has *conceptual equivalence*, and this obviously goes far beyond the equivalence implied by such a phrase as response generalization (cf Kendler, Ch 5)

Measuring generalizability (or transferability) would seem to be something which the experimenter on problem solving sometimes does (e g, Katona, 1940) It is my belief that the inference of problem solving, as a process distinct from other types of learning, *requires* the demonstration of generalizability, and therefore such a demonstration should be made much more frequently than has been the case in the past

It seems clear that the problem solving investigator intends that his learning situation (when successful) shall establish a capability for the *S* to solve a *class* of task which he defines, and not simply a single member of that class Almost any example could be chosen to illustrate this point —let us take Maier's Candle Problem (Maier, 1933) There are several different learning situations which *could* establish the criterion performance of fitting together tubes and clamps so as to blow out a candle 8 feet away One of the most obvious would be to have the *S* watch this being done, or even look at a picture showing the final construction It would

surely be agreed that such a learning situation would be quite effective in bringing about the change in performance which the experimenter has decided to be interested in But this is what some writers have called 'reproductive' learning, and it is not the inference which the experimenter in this case *wants* to make—instead, he wants to say that problem solving has occurred It would appear that the only criterion which can make such an inference certain, is to show that the performance is *immediately generalizable to an entire class of tasks* of a type which the experimenter defines (or preferably, has defined beforehand) For this example, such a class might be 'using clamps to hold short tubes together so as to make long tubes' As is true for certain other investigators, Maier does not actually make this demonstration of transferability Yet there seems to be little doubt that he believes it could readily be made, in fact, the reader usually has little doubt that it could be made

The requirement to establish generalizability as a condition which justifies the inference of problem solving, as opposed to 'reproduction," has been recognized implicitly by some investigators, explicitly by others The necessity is, for example, fairly obvious in the case of monkeys solving oddity problems (Harlow, 1949) No one would be particularly excited to learn that a monkey had solved a single oddity problem Rather, it is the fact that the monkey solves a variety of problems with a number of different stimulus objects, resembling each other only insofar as they can be classified as 'oddity problems" by the experimenter, which makes it possible to infer problem solving (or learning set) This appears to be the basis of the argument made by Wertheimer (1945) regarding "good' solutions for the area of a parallelogram Wertheimer wishes that children would acquire a capability of 'finding areas for four sided figures," which is immediately generalizable to all four sided figures, rather than learning merely to reproduce a formula ($b \times h$) for the area of a rectangle In other words, his criterion for a 'good' solution is generalizability within a class of tasks Transfer among tasks is used to indicate the influence of set in the solution of water jar problems (Luchins, 1946) No writer, though, has given this factor a more thorough emphasis than Katona (1940) When *Ss* solve his card trick or matchstick problems nonreproductively, he expects them to be able immediately to solve other problems of the same class Not only does he apply this criterion of transferability in his experiments but he recognizes the critical role that transfer plays in his distinction between organization" learning and "reproductive learning

The demonstration of immediate generalizability in problem solving carries the implication that the process which underlies successful performance is not "response tied" Not only may the solution to a single prob-

lem have a number of equivalent forms, as we have already seen, but the capability which is acquired by the *S* must be one which underlies a great variety of specific performances all appropriate to a single class of task. The point is, if monkeys have learned to solve "oddity problems," one assumes that this solution could be exhibited in almost any response mode, whether by stepping on platforms, pushing buttons or opening doors. Similarly, if a human *S* has solved the problem called "adding fractions," one expects this could be exhibited in any of an infinite number of specific responses. It would seem, therefore, that if one wishes to speak of "correct responses" in problem solving, as opposed to "correct answers," he must use the former phrase in a purely metaphoric sense. It would perhaps be preferable to avoid using it, for this reason. The property of generalizability of problem solution sets this form of learning in marked contrast to serial, paired-associate, and discrimination learning. In such forms, the finding of immediate transferability would presumably be disturbing to the experimenter, in problem solving, it is confirming. Accordingly, the demonstration of generalizability may be considered a good differential criterion for the distinguishing of problem solving from other varieties of learning.

Independent Variables

What are some of the independent variables which may be presumed, or shown, to affect problem solving performance? What are the various parts of the "learning situation" that the experimenter either controls or manipulates? Two major classes of events come immediately to mind when one asks these questions. First, there is the presentation of stimuli or "stimulus objects," like the string and weights in the pendulum problem, the pattern of matches in a matchstick problem, the panel of large and small squares in a reversal shift problem. Second, there are instructions to the *S*, given so that he will "understand the problem," or perhaps for the deliberate purpose of establishing a "set." Whatever the rationale, it seems likely that instructions in problem solving may serve functions other than simply restricting the *S*'s behavior to a form desired by the experimenter for measurement purposes. Consequently, we need to examine them more closely than would otherwise be the case.

The Stimulus Situation

One can identify certain stimuli or objects in the stimulus situation which retain their physical identity from the beginning of problem solving to its end. For example, the sticks and clamps in Maier's hatrack problem are a part of the original stimulus situation, and remain the same, physi-

cally speaking, when the solution has been achieved Similarly, in matchstick problems, the matches retain their physical identity during the problem presentation as well as in the solution But the identifying of such unchanging status for stimulus objects may be a somewhat trivial thing to do For the most striking set of events in problem solving, most authors agree, is a change in "patterning' of the situation from the beginning to the end of the session Duncker (1945), for example, conceives of problem solving as a set of events which get the learner from a "given situation" to a "desired situation" One can say also that problem solution is not bound up with the physical identity of the matchsticks before and after solution, but rather with the *change in pattern* which they assume after solution as compared with before

Several writers on problem solving have also emphasized that the subject *adds* something to the stimulus situation when he solves the problem This appears to be the meaning of phrases like "going beyond the information given," used by Bruner (1957), and 'filling in gaps," as discussed by Bartlett (1958) It may be that these expressions are simply other ways of expressing the idea of 'reorganization" (or a similar word) mentioned by other writers (e g, Duncker, 1945, Katona, 1940, Maier, 1945) The importance of these ideas for a description of independent variables, however, is this the stimulus situation for problem solving must be one which is capable of being specified in conceptual categories, rather than in terms of physical stimulus characteristics This is true because whatever 'fills the gap' must be an abstraction, rather than a thing with particular physical characteristics, whatever is "newly organized' must be a conceptual entity The number series 1 2 5 — 17 33 may be filled with the *number* 9, but if it is, this means that the original stimulus situation has been composed of *numbers*' not simply of numerals on a page

Accordingly, it may be said that attempts to specify the stimulus situation for problem solving in terms of physical stimulus attributes will be quite inadequate The idea of making this attempt has apparently not even occurred to most investigators of human problem solving but if we are to allow animals to be *Ss* in such experiments, it may do so The single example of the monkeys with oddity problems should clearly demonstrate the reason So long as an animal is making responses which are 'stimulus tied ' that is, correlated with a particular physical configuration his behavior can be accounted for in whatever terms are used to explain discrimination learning But as soon as his responses are found to be correlated not with a particular stimulus but with a stimulus *class* which the experimenter names (whether "large," "middle,' "odd," 'below, or whatever), a fundamental requirement for problem solving

behavior has been met [Excluded from consideration are those instances of "near-transposition" which can presumably be attributed to the effects of primary stimulus generalization (cf Kuenne, 1946)] It would seem evident, then, that the specification of a stimulus situation for problem solving must be in terms of abstracted properties of stimuli or their physical arrangements Such a state of affairs places problem solving in contrast to trial-and error, discrimination, or even to human verbal reproductive learning, for in all of these instances the experimenter strives to describe stimulation in physical terms, and to maintain physical identity from trial to trial or from session to session At the very least, even when physical variation is permitted, he attempts to stay within the bounds set by stimulus generalization For problem solving, certain kinds of physical variation are unimportant because they are "coded" in the same way (like saying "2" or "two" for the number 2), while others may be introduced deliberately (in order to prevent the acquiring of specific habits) What remains constant about the stimuli is the class to which they belong, a conceptual property

We seem to need a convenient way of making this distinction between stimuli which are intended to have a direct relationship with responses, and stimuli which are not so intended, but for which an abstract property may be related to responses The first kind of stimuli may be said to function as *definite identities,* that is, one looks for a relationship between a particular stimulus and a performance to be learned The second kind, in which the class of stimuli is related to performance, may be called *indefinite identities*[2] For the study of problem solving, one needs to specify the stimulus situation in terms of indefinite identities

Instructions

Almost all experiments on human learning utilize instructions They are used for various purposes—to motivate the *S*, to give him an achievement set, to provide for his adjustment to the experimental situation Besides these things, they permit the experimenter to put constraints on the *S*'s behavior—by getting him, for example, to make oral responses which are audible, to watch the window in a memory drum, to push one set of buttons but not others All of these things, and perhaps others, are effects of instructions that are common to learning studies We may not understand very well how they work, but we satisfy ourselves that they are in fact capable of removing large chunks of unwanted variance from the experimental situation These functions need not be considered further here

[2] I am indebted to my colleague Myron Goldstein for these phrases distinguishing two types of stimuli

In problem solving situations, instructions seem to go considerably farther than would be expected by these common functions Investigators of problem solving sometimes state that instructions must "define the problem" for the *S*, that they must give him an "understanding of the goal', or perhaps even that they may "establish a set" or "introduce direction" Is all this necessary? Have experimenters in this field become somewhat soft headed and perhaps careless, in allowing instruction to do so many things in problem solving? How can we control a learning situation containing such relatively ill-specified variables?

In problem solving studies, instructions appear to have, besides the *constraining* function previously mentioned, a *contributing* function Although this may occur in other types of learning situations, it is given major emphasis in problem solving Instructions in the latter case stimulate the learner to engage in activities which presumably contribute to the attainment of problem solution, although these activities are not necessarily identifiable in the final performance It is frequently the case that instructions in a problem solving experiment embody some of the most important independent variables These same variables may, of course, be manipulated in other ways when one is dealing with animals, and they may sometimes be incorporated in "pretraining" sessions with human beings But most typically, with human *S*s, they are handled by statements that carry communication from the experimenter to the *S* (By "instruction" is meant any communication to the *S* which can be shown to alter his behavior, generally speaking, this is everything else except the stimulus objects considered in the previous section In some experiments, to be sure, part of the instructions may be printed rather than spoken) It will be instructive to consider what difference they may be expected to make in the solving of problems

If the functions of instructions are to be described, we need an example It seems worthwhile to quote a set of instructions which has been used by several investigators, but originally designed by Maier (1930) They are as follows 'Your problem is to construct two pendulums, one of which will swing over this point (cross indicated on floor, see Figure 2) and one which will swing over this point (other cross indicated) These pendulums should be so constructed that they will have a piece of chalk fastened to them which will make a mark (which can be seen) on the points on the floor just indicated Naturally you must have something to hang the pendulums to That is for you to worry about Don't try to move the table about Otherwise do anything you want to This material is at your disposal That chair, however, is not to be part of your construction, you may use it for a work bench or a place for meditation, or anything you wish so long as it is free when you are through Ask any question you

wish I'll be glad to assist you in building, only you must tell me what to do" (Maier, 1930, p. 118).

Besides these instructions given to all groups in his experiment, Maier used an additional set with some groups, which he describes as dealing with "parts" of the problem. In brief, these instructions were as follows: Part A, how to make a plumb line by using a clamp, a pencil, and string; Part B, how to make a long pole out of two short ones, using a clamp; Part C, how to hold up a screen (for a projector) with two poles, by wedging.

Still greater variety was given to the "instructions" variable by Maier by including for one group the strong suggestion that what must be done is a matter of "combining ideas", and for still another group, a set of additional instructions called *direction*. The latter went as follows:

"I should like to have you appreciate how simple this problem would be if we could just hang the pendulum from a nail in the ceiling. Of course, that is not a possible solution but I just want you to appreciate how simple the problem would be if that were possible. Now that it is not possible the problem is, as you may find, really quite difficult" (Maier, 1930, p. 119).

With this great variety, there should be little doubt that the major independent variables in problem solving were considered by Maier to be contained in one or more forms of these instructions. With these as an example, I should like now to try to identify what functions instructions were serving in Maier's experiment, with the hope that there will be some generality to this determination.

Each of the following functions will be expressed as something that the *S* was (hopefully) able to do, after hearing Maier's instructions and responding to them.

1. *Identify the terminal performance required.* Maier's initial instructions are almost entirely concerned with this function. They tell the *S* that what is required is a construction which will contain two pendulums able to mark the floor but which will not contain the chair.

2. *Identify parts of the stimulus situation.* While not devoting a great deal of time to it, Maier's instructions do try to ensure that the *S* can identify a pendulum (as something that swings freely weighted on one end), a chalk mark, a clamp. Of course, the problem was one in which these items of the stimulus situation were chosen to be familiar; accordingly, identification was assumed to be already accomplished for most subjects.

3. *Recall relevant subordinate performance capabilities.* This seems to have been a prominent and probably also very important function of Maier's instructions. The first of these contained in the initial instructions may be identified as "how to suspend a pendulum by hanging." The others are identified as Parts A, B, and C, described in a previous paragraph.

4. *Channeling thinking.* The instructions appear to have the function of guiding

the *S*'s thinking into profitable channels In other words they serve to weaken or eliminate certain hypotheses that the *S* might reinstate and to strengthen others The word "construction" in the initial instructions may have this (additional) function The suggestion "put these ideas together" definitely seems to fall in this category And of course, so do the "direction" instructions which have the tendency to strengthen hypotheses concerned with "hanging the pendulum from the ceiling"

Let us now consider whether there are other sources of evidence for these variables, in other problem solving situations, and perhaps also, the extent to which they can be manipulated within an experimental setting

Identifying the terminal performance—This is a variable which is not often dealt with in any way except to ensure its presence While one can conceive of getting this variable to vary in a continuous fashion (e g, varying the degree of precision with which the *S* can identify the required performance), the most obvious variation is in or out If instructions are not included which enable the *S* to identify terminal performance, presumably the probability of his attaining a solution will be reduced There is not much formal evidence to be found for this prediction, however, because most investigators are inclined to emphasize the importance of "defining the goal" In the course of a study by Gagné and Brown (1961) on the problem of stating formulas for the sum of *n* terms in number series, we noted in our initial observations a number of *S*s who responded to the problem with a numerical answer, rather than with a symbolic one This answer was, of course, wrong, and would under these circumstances be counted as an unsuccessful solution This observation led us to revise our instructions (contained in a learning program) to ensure that the general configuration of a correct solution could be identified by the *S* It is possible that this kind of observation has led other investigators of problem solving to carefully include this variable in their instructions, rather than leave it out At any rate, its presence in instructions is easy to identify, and is virtually universal

Identification of aspects of the stimulus situation—The problem solver must respond correctly to objects in the stimulus situation, particularly in the sense of differentiating among them If the word "pendulum" is used presumably the problem cannot be solved unless a correct identification can be made of this word in the sense of being able to pick out pendulums from constructions which are not pendulums In symbolic problems such as those in mathematics identifications must be made of the symbols, a coefficient in an algebraic expression must be differentiated from an exponent Achieving solution may be considered to be impossible if such identifications are incorrect Although this variable in the learning

situation has not been systematically studied, frequent reference is made to it in practical situations Studies of electronic troubleshooting frequently emphasize the importance of "perceiving the situation correctly,' which surely conveys the meaning of concept identification and differentiation

One can most readily conceive of varying this factor of stimulus-object identification in an all-or none fashion, as is true for identification of the terminal performance When the problem contains a number of objects of the sort which tend to be confusable, it would presumably be possible to vary the "amount of learning"—i e , the number of identifications acquired or to be acquired It may be supposed, though, that the variable truly operates in all-or none fashion in being prerequisite to problem solving, if any of the required concept identifications cannot be made, no problems involving these identifications can be solved, regardless of the operation of other variables

Recall relevant capabilities—As has been stated, this factor is given considerable prominence as a function of the set of conditions that are here called "instructions" The experimenter is concerned to establish high recallability of "part activities," which I suggest might be called *subordinate capabilities* Maier's "parts" fit neatly into this category Moreover, the work of Weaver and Madden (1949) and of Saugstad (1955) show clearly that increasing the recallability ("availability') of subordinate capabilities leads to significant increases in the proportion of successful solutions to Maier's problems Katona's results (1940) indicate that when an arithmetic or spatial principle is made highly recallable, the proportion of solution of matchstick problems rises In simpler problem situations Kendler and Kendler (1956) establish high recallability of subordinate activities in children, in order to study the solution of problems requiring an integration of these activities In the case of anagram solution (Maltzman, Eisman, Brooks & Smith, 1956, Rees & Israel 1935) the establishment of high recallability of certain word categories which can be used in problem solution accomplishes striking changes in number of solutions

Recall of subordinate capabilities is effected in several specific ways in the problem-solving experiment The *S* may simply be "reminded as he is, for example, when instructions state a word category in anagram problems Or, he may actually be required to practice the subordinate activities in one or more trials, as when he is asked to state the functions of stimulus objects (Saugstad, 1957) There appears to be a continuity between those situations which accomplish the required recall by means of instructions per se, and those which bring recall about by pretraining as in anagram studies, the Kendler studies with children's inferences and studies of preavailability of concepts (Judson Cofer, & Gelfand 1956 Saugstad 1955,

Staats, 1957) When only instructions are used, one can consider that recall is being effected by the method of recognition, whereas the method of recall (reinstatement) may be employed either as part of the instructions or in a separate training period

It is clear that what is being recalled in the problem solving situation is not the criterion performance, but some kinds of subordinate capabilities The most obviously manipulable variable used for this purpose is number of repetitions However, one should probably not lose sight of other factors affecting recall, particularly those of timing It is not unreasonable to suppose that the passage of time following the initial establishment of subordinate capabilities may affect their recall In addition, it may be necessary that some critical time interval be ensured between the recall of more than one subordinate capability, in order for problem solving to take place It is entirely conceivable that the processes of sort term memory (Postman, Ch 4) may be directly involved in these events Underwood s remarks (1952) about contiguity in responses as a factor in concept learning appear to make a similar, or at least a related, point

Channeling thinking—Besides the 'direction' used by Maier, several other procedures have been followed in channeling the *S* s thinking along certain lines and away from others Katona (1940) used certain types of demonstrations both in his card trick and matchstick problems, and found them to have considerable effectiveness, as was also found by other investigators (Hilgard, Irvine, & Whipple, 1953) A number of studies have employed 'helps" or 'hints," which also appear to be performing this function The notion that certain habit family hierarchies may be increased in strength relative to others by certain elements contained in instructions occurs in Maltzman's (1955) theory of problem solving Of course, amount of guidance is a variable that may take a zero value, and Saugstad s (1955) study clearly shows that even with Maier's problem, direction may be unnecessary for problem solving if subordinate capabilities are adequately recalled

Little more can be said about this factor at the present time Few investigations of the guidance of thinking have been made in recent years It is even possible that it is entirely an unnecessary variable in problem solving or that it is a misnamed one But in view of the prominent part it has played in several problem solving studies, it would probably be unwise to ignore it entirely It seems likely that there may be a set of manipulations of instructions that serve to strengthen certain "channels of thought" and to weaken others, serving thereby to speed up the process of problem solving But we do not know clearly what these manipulations are, or how to measure them

Individual Differences Variables

Investigators of human learning are always beset by problems of how to handle individual differences in an experimental setting Even in studies of conditioning, the problem of "voluntary" eyelid responses plagues the experimenter In the learning of paired associates or serial lists, the very large differences among individuals in such measures as number of trials to learn gives rise to a variety of techniques for obtaining group measures, perhaps none of which is considered fully satisfactory Concept learning studies must contend with individual differences that are even more diverse Consequently, it is not surprising that the prominence of individual differences in studies of problem solving attains the status of a crucial problem, rather than simply being a vexatious nuisance

Generally speaking, the problem of individual differences is handled, so far as the final performance is concerned, by abandoning the attempt to measure "degree of learning" and by counting the number of individuals who achieve or do not achieve the required solution It would be wrong to level the accusation that by adopting this procedure, such studies have become nonquantitative The proper quantification of problem achievement may in fact be one of pass or fail, as has previously been stated

When viewed from the side of the independent variable, individual differences in problem solving studies can be handled as they are in investigations of transfer of training The instructional situation is often designed to produce differences (among contrasted individuals) in those subject variables previously described, namely, in identification of terminal performance, in identification of stimulus-objects, or in recall of subordinate capabilities Having established conditions designed to bring about such differences, the individuals can then be subjected to a standard situation (including the stimulus situation and other instructions), and performance on the final task observed Were individual differences treated this way systematically in problem solving studies, the tightness of design would presumably be comparable to those in transfer of training studies

But although the potentiality for dealing directly with individual differences as manipulable independent variables is clearly present, it must be said that this has not often been done If they are to be dealt with in this way, individual differences must be *independently measured*, not simply assumed to result from some changes in the instructional situation If one has a serious interest in determining the effects of individual differences he must measure independently whether or not the individuals *can* identify the terminal performance, or *can* identify relevant aspects of the stimulus situa-

tion or *do* possess certain relevant capabilities The importance of this approach is suggested particularly by the various findings on the effects of variations in subordinate capabilities on problem solving achievement (Katona 1940, Maier, 1930, Saugstad, 1955) In some of these studies the effects of instructions in producing individual differences is only assumed But in Saugstad s study the possession of a subordinate capability of using a glass tube to conduct an air stream was measured directly The implication is that problem solving needs to be studied as a two-stage process, as one studies transfer of training The first stage involves the measurement of subordinate capabilities, the second stage involves the measurement of the final performance

The Nature of Problem Solving

As a means of summarizing prior to considering the relationship be tween problem solving and other forms of human learning it would seem desirable at this point to try to weave a design which pulls together the threads of our previous discussions What is the set of events that is called problem solving?

The individual learner, we assume, comes to the problem solving situa tion with certain basic aptitudes But even more important, he comes to it with certain previously *learned* capabilities which play a very direct role in his problem solving behavior, in the sense that their presence or absence determines whether or not he will successfully achieve solution The experimenter exposes to the learner a stimulus situation containing stimuli which first must be identified in a conceptual sense Some of these *indefinite identities* may have been previously acquired while some may be new The experimenter communicates with the learner by means of instruc tions These change the behavior of the *S* in four ways identifying new stimuli (i e establishing new indefinite identities) identifying the expected form of the terminal performance recalling previously acquired capabili ties and channelling thinking in a relevant direction These four factors may be made to assume certain values or varieties for experimental pur poses The individual then engages in problem solving requiring a certain amount of time At the end of a reasonable period the experimenter ob- serves that the *S* either has or has not solved the problem By this is meant that he has or has not provided an answer to a class of tasks that define the problem (for the experimenter) Unless something has gone wrong the achievement of solution to one member of this class of tasks should mean immediate generalizability to any other member of the class

In this account certain parts have a ring of familiarity and seem to be similar to those encountered in other forms of learning At the same

time, there is a striking difference which can perhaps best be summarized in this way: the learning situation for problem solving never includes performances which could, by simple summation, constitute the criterion performance. In conditioning and trial and error learning, the performance finally exhibited (blinking an eye, or tracing a path) occurs as a part of the learning situation. In verbal learning, the syllables or words to be learned are included in the learning situation. In concept learning, however, this is not always so, and there is consequently a resemblance to problem solving in this respect. Although mediation experiments may present a concept during learning which is later a part of the criterion performance, many concept learning experiments do not use this procedure. Instead they require the *S* to respond with a performance scored in a way which was not directly given in learning (the stating of an abstraction such as "round" or "long and rectangular"). Similarly, the "solution" of the problem is not presented within the learning situation for problem solving. Concept formation and problem solving are *nonreproductive* types of learning.

An implication of this difference is that in problem solving there is a clear distinction between the learning situation and the measurement of the final performance. When preparing to study problem solving, one first decides upon a terminal performance to be measured. Having done this, he then takes the entirely separate step of preparing a learning situation (or "learning program") which cannot, by the rules of the game, include the specific final task later to be employed for measurement. What it does include, as previously suggested, is (1) a stimulus situation and (2) a set of instructions having the functions previously identified.

A corollary implication is that problem solving is by its nature a transfer of training situation. The paradigm which fits is: first, differential learning treatment; second, measurement of performance on a standard class of tasks. Of course, various refinements of this paradigm may be required in specific experiments for purposes of control. But besides this, the suggestion is that the underlying *process* of problem solving, whatever it turns out to be, must have a lot to do with transfer of training. And one is tempted to remark, how little we know about transfer of training!

Having noted this basic distinction, let us turn to a consideration of how problem solving may be related to other (possibly simpler?) types of learning. Table 1 contains paradigms and descriptions of various forms of learning, from conditioned responses to problem solving. Basically, its conception is that what we call learning ranges from the relatively simple response learning to complex problem solving. Increasing complexity is seen to reside not so much in what is learned, as in the nature of what has to be *pre-available* (either shown to be or assumed to be) (cf. Duncan, 1959) in order for various types of learning to occur. Thus verbal paired associate

learning in its pure form occurs when the responses (or response-connections) are already available (cf. Underwood and Schulz, 1960), originally made so, presumably, by previous learning. Concept learning is in turn based upon the assumption, or actual establisbment, of preavailable verbal "labels," which have previously been acquired as paired associates are. (This prior process is, of course, the stimulus coding discussed by Underwood, Ch. 2.) And so on until the most complex form, problem solving, is reached, wbich depends upon the preavailability of capabilities acquired in all the other forms of learning.

Let it be clear that I shonld not like to claim tbat the table is an exhaustive one, or even that all the subtleties of the various forms, as they

TABLE 1

A SUGGESTED ORDERING OF THE TYPES OF HUMAN LEARNING

Type	Paradigm[a]	Description
Response learning	S-R	Establishment of a response-connection to a stimulus specified along physical dimensions
Chaining	S-R ~ S-R	Establishment of chains of response-connections.
Verbal learning (paired-associates)	S-r ~ [s-R]	Establishment of labeling responses to stimuli varying physically within limits of primary stimulus generalization. Previous "response learning" assumed (as indicated by brackets).
Concept learning	S r~~s S-r~~s → Concept — S-r~~s	Establishment of mediating response to stimuli which differ from each other physically ("classifying").
Principle learning	—Concept → Rule — —Concept	Establishment of a process which functions like a rule "If A, then B," where A and B are concepts.
Problem solving	—Rule → Higher-order rule — —Rule	Establishment of a process which "combines" two or more previously learned rules in a "higher-order rule."

[a] The paradigms shown have been designed to depict what is learned, and not the learning situation which leads to this result. In addition, it may be noted that beginning with concept learning, only the central portions of the inferred chains are shown.

may occur, for example, in the varieties of conditioning discussed by Grant and Kimble (Ch 1) are validly represented The main purpose is to present the idea that the more complex forms of learning depend upon processes which have been previously acquired in the simpler forms, in a hierarchical fashion The suggestion is, therefore, as one proceeds to observe more and more complex forms of learning, he is led to search for variables which are farther and farther removed from the stimulus situation, and which tend increasingly to exist as previously acquired capabilities of the *S* This reflects again our previous argument that problem solving is a matter of transfer of training, and also that some of its most important independent variables are to be found in the capabilities of the individual

Let us look a little more closely at some of the complex forms of learning shown in the table Concept learning is presented here as a form in which a number of physically different stimuli are put into a single class by the *S* Usually, it is assumed that he has previously acquired a variety of verbal labels for the stimulus objects which are presented They can be called "round," "square," "tall," 'short", any of a variety of verbal responses can be made to them, each of which has at some time or other been acquired as a "verbal association" Concept achievement is observed when the *S* becomes capable of responding to these different physical objects as if he were placing them in one or more classes, like "round," "square and large" The conditions of establishment of such concepts, as well as the properties ascribed to them by various investigators, are discussed by Kendler in Chapter 5

As the table shows, a form of learning can be identified which appears to be more complex than concept learning, and yet simpler than some of the problem solving we have talked about This may be called "principle learning," or "rule learning" Presumably, it falls within the scope of this paper Yet it does not appear to have been extensively studied in a form which fits the paradigm shown [The study of Shepard, Hovland, and Jenkins (1961) presented a task which may be considered a somewhat curtailed representation of a principle, as will be clear from the following discussion] It would not be difficult, though, to imagine an experimental situation in which it could be investigated First of all, the pre-availability of concepts would be assumed or measured Suppose the individual has learned to classify an array of stimuli as "round' or "square," and another array as "the odd one" or "not the odd one" What would be investigated in this situation are the conditions leading to the use of the principle "When round, choose the odd one" The stimulus situation might consist of presentations of the two arrays of stimuli (in suitable variety of physical dimensions) simultaneously Undoubtedly, one could simply present such a situation in repeated trials, and observe the errors

made or the number of trials required to attain the rule But this is not the only kind of learning situation one thinks of, particularly if the question 'how could such a principle be most rapidly achieved' is of interest The idea of manipulating instructions to establish "sets," to channel thinking, and to bring about recall of the previously acquired relevant concepts, is one which might be prominently exploited experimentally In a sense, then, *principle learning* appears to be a simple form of what has been called problem solving

The more complex form is exhibited in the last line of the table This begins with the assumption (or the establishment) of pre available rules It ends with the demonstration that the individual has acquired a new 'higher-order rule," the capability of solving a new kind of problem Let us consider an example which is as simple as possible A beginning student of algebra is set the problem "Multiply X^2 and X^3," which he has not seen before at all First of all, it is clear that he must previously have acquired certain *concepts*, namely, the concepts of variable and exponent (in the sense that he must be able to recognize these stimulus objects as "variables with exponents") Second, he must previously have acquired some *rules* One is a rule for multiplication, approximately stated as 'multiplying a number by *n* means adding the number *n* times" Another is a rule for exponents, roughly, "an exponent *r* means multiplying the number by itself *r* times" Can the student solve this problem? Perhaps he will need to be reminded that he knows these subordinate rules Perhaps he will need some *guidance* or "direction" for his thinking Whatever may be required to achieve it, what is wanted in solution is the discovery of a new and more inclusive rule to the effect "Multiplying identical variables with exponents is done by multiplying the variable by itself the number of times represented by the *sum* of the two exponents" The attainment of such a higher-order rule can, under proper circumstances, be inferred by the correct answer X^5, as well as by correct answers to any and all other tasks belonging to the class multiplying variables with exponents"

Thus it appears that the investigation of problem solution in a situation like this, as well as in many more complex ones (involving greater numbers of rules and concepts) may be reduced to an experimental situation in which the independent variables fall into the two broad categories of *subject capability* and *instructions* The capabilities that the *S* begins with include previously acquired rules previously identified concepts, which in turn have been based upon previously established labeling, and ultimately also on previously learned response-connections If one desired to study the relatively pure "final stage" of problem solving, he would presumably begin by measuring these capabilities, and then proceed by manipulating the variables contained in instructions

Methodological Implications

It is possible that the implications of this discussion for methodology are already clear. But I shall try to summarize them here.

From the standpoint of the analysis of problem solving just previously presented, most studies of problem solving have a mixed appearance. Some of them try to relate problem solution directly to variables in the stimulus situation. Others, a little more sophisticated perhaps, attempt to study the effects of concept mediation. The point of view proposed here is that such studies can only yield ambiguous results, because they are not able to bring the most important independent variables under control. Such variables are the individual capabilities of the *S*, conceived as previously acquired "rules," whose availability may readily be measured independently of problem solution itself. Such variables may also be viewed as constituting the largest sources of the variance often referred to as individual differences existing before the experiment (cf. Gagné & Paradise, 1961). Until these sources of behavioral change are brought under control, or varied systematically, differences among *S*s will continue to be a difficult and perplexing problem for the experimenter interested in the causes of problem solving.

To study the phenomenon of problem solving, an experimenter needs first to define a class of criterion tasks to be used for final measurement. Such tasks should require the use of a common principle in their solution, preferably a principle which can be clearly related to some subordinate principles, achievement of which can likewise be independently measured. Having done this, the experimenter may decide he is interested in this terminal stage of the process of problem solution. In such a case, he undertakes first to ensure that the subordinate principles are in fact available to the individual *S*s he is studying. Then he is ready to design an experiment testing the effects of one or more aspects of the "instructions" on the attainment of problem solution. As previously suggested, the most obvious variables here are those pertaining to (1) identification of stimulus objects and of terminal performance, (2) recallability of subordinate principles, and (3) guidance of thinking. These are perhaps not as well defined variables as one would hope they will eventually become. And there may very well be others.

This is not the only route which can be followed. The experimenter may decide that he is interested, not simply in how the availability of rules affects problem solving, but in the contribution made by concepts, verbal associates, or even fundamental responses. According to the view here presented, such interests should logically lead to experiments which study a kind of "progressive transfer of training." While perhaps a little less

clean-cut, this seems to be a perfectly feasible kind of experiment. One would first independently measure, say, the establishment of concepts, then proceed to the establishment of rules, and finally to the achievement of a higher-level rule in problem solving. Each successive measure would be conceived as indicating transfer from the subordinate capabilities achieved in the previous learning stage.

The conclusion to which we are led is that problem solving has a very definite relationship to other forms of learning. If problem solving is to occur, it must have been preceded by a variety of simpler forms of learning that are more directly "tied" to the stimulus situation and its variables. If one has the aim of studying the phenomenon of problem solving, it is desirable, even necessary, to demonstrate that the capabilities established by these other learning situations are pre-available. Once this has been done, it may be expected that the factors which affect the attainment of problem solution will be found in that part of the experimental situation usually called "instructions."

References

Bartlett, F C *Thinking An experimental and social study*. New York Basic Books, 1958

Bruner, J. S, Goodnow, J J, & Austin, G A *A Study of thinking* New York: Wiley, 1956

Bruner, J S Going beyond the information given In *Contemporary approaches to cognition* Cambridge Harvard University Press, 1957.

Duncan, C. P. Recent research on human problem solving *Psychol. Bull*, 1959, 56, 397–429

Duncker, K On problem solving *Psychol Monogr*, 1945, 58, Whole No 270

Ewert, P. H, & Lambert, J F Part II The effect of verbal instructions upon the formation of a concept *J Gen Psychol*, 1932, 6, 400–413

Gagné, R. M., & Brown, L. T Some factors in the programming of conceptual learning *J exp Psychol*, 1961, 62, 313–321

Gagné, R. M., & Paradise, N E. Abilities and learning sets in knowledge acquisition *Psychol Monogr.*, 1961, 75, Whole No 518

Gagné, R M, & Smith, E C, Jr. A study of the effect of verbalization on problem solving. *J exp Psychol*, 1962, 63, 12–18

Harlow, H F The formation of learning sets *Psychol Rev*, 1949, 56, 51–65

Highland, R. W, Newman, S E., & Waller, H S A descriptive study of electronic trouble shooting. In G Finch and F Cameron (Eds), *Symposium on Air Force Human Engineering, Personnel and Training Research* Washington, D C: National Research Council—National Academy of Sciences, Publ. No 455, 1956 pp 173–198

Hilgard, F R, Irvine, R. P, & Whipple, J E Rote memorization, understanding, and transfer An extension of Katona's card trick experiments. *J exp Psychol*, 1953 46, 288–292

John, E R Contributions to the study of the problem solving process *Psychol Monogr* 1957 71, Whole No 447

JUDSON, A J, COFER, C N, & GELFAND, S Reasoning as an associative process II Direction" in problem solving as a function of prior reinforcement of relevant responses *Psychol Rep*, 1956, 2, 501–507

KATONA, G *Organizing and memorizing* New York Columbia University Press 1940

KENDLER, H H, & KENDLER, T S Inferential behavior in preschool children *J exp Psychol*, 1956, 51, 311–314

KENDLER, H H, & KENDLER, T S Effect of verbalization on reversal shifts in children *Science*, 1961, 134, 1619 1620

KENDLER, T S, & KENDLER, H H Reversal and nonreversal shifts in kindergarten children *J exp Psychol*, 1959, 58, 56–60

KUENNE, M R Experimental investigation of the relation of language to transposition behavior in young children *J exp Psychol* 1946 36, 471–490

LING, B C The solving of problem situations by the preschool child *J genet Psychol*, 1946, 68, 3–28

LUCHINS, A S Classroom experiments on mental set *Amer J Psychol*, 1946, 59, 295–298

MAIER, N R F Reasoning in humans I On direction *J comp Psychol*, 1930, 10, 115–143

MAIER, N R F An aspect of human reasoning *Brit J Psychol*, 1933 24, 144–155

MAIER, N R F Reasoning in humans III The mechanisms of equivalent stimuli and reasoning *J exp Psychol*, 1945, 35, 349–360

MALTZMAN, I Thinking From a behavioristic point of view *Psychol Rev*, 1955, 62, 275–286

MALTZMAN, I, EISMAN, E, BROOKS, L O, & SMITH, W M Task instructions for anagrams following different task instructions and training *J exp Psychol*, 1956, 51, 418–420

PETERSON, J, & LANIER, L H Studies in the comparative abilities of whites and negroes *Ment Measmt Monogr* 1929, Serial No 5

RAY, W S Complex tasks for use in human problem solving *Psychol Bull*, 1955 52, 134–149

REES, H, & ISRAEL, H An investigation of the establishment and operation of mental sets *Psychol Monogr*, 1935, 46, Whole No 210, 1–26

REID, J W An experimental study of analysis of the goal" in problem solving *J gen Psychol*, 1951, 44, 51–69

RUGER, H The psychology of efficiency *Arch Psychol* N Y, 1910 Whole No 15

SAUGSTAD, P Problem solving as dependent upon availability of functions *Brit J Psychol*, 1955, 46, 191–198

SAUGSTAD, P An analysis of Maier s pendulum problem *J exp Psychol* 1957, 54, 168–179

SHEPARD, R N, HOVLAND, C I., & JENKINS, H M Learning and memorization of classifications *Psychol Monogr* 1961, 75, Whole No 517

STAATS, A W Verbal and instrumental response hierarchies and their relationship to problem solving *Amer J Psychol*, 1957, 70, 442–446

UNDERWOOD B J An orientation for research on thinking *Psychol Rev.*, 1952, 59, 209–220

UNDERWOOD, B J, & SCHULZ, R W *Meaningfulness and verbal learning* Philadelphia Lippincott, 1960

WEAVER, H E, & MADDEN E H "Direction" in problem solving *J Psychol.*, 1949, 27, 331–345

WERTHEIMER, M *Productive thinking* New York Harper 1945

Learning and Problem Solving.

COMMENTS ON PROFESSOR GAGNE'S PAPER[1]

TRACY S KENDLER
Barnard College

Professor Gagne has had a difficult assignment with which he has done so well that I cannot agree to differ So I will limit my discussion to a few remarks intended to help clarify some of the issues that must be faced in this research area

Science consists of continuous explorations of new and hitherto mysterious phenomena in order to render then familiar and understandable One of the devices in common use is to relate the unexplained events to some process which is better understood Although, in our more optimistic moments, it may seem that some progress is evident, we are nevertheless likely to agree that problem solving is still, largely, an engrossing mystery Perhaps we can also agree that, although we are far from finished with the task of making learning understandable, we have certainly made it familiar As a consequence there are quite a few psychologists who, like Gagne, have bravely taken up the tools, the language, and the theories of the relatively familiar field of learning research to apply to the relatively unfamiliar field of problem solving research

THE RELATION OF LEARNING AND PROBLEM SOLVING

The attempt to relate learning and problem solving may seem such an eminently sensible thing to do, that it bears little discussion The ultimate justification for reconciling problem solving with learning depends upon whether the research and theorizing it produces will make either process better understood And this, of course, is an empirical issue But it is an issue about which we who are concerned with understanding problem solving behavior, are impelled to make predictions, in order to guide our research strategy Explicitly or implicity, we make some important assumptions It is better to deliberately examine these assumptions than to leave them buried but fundamental

[1] The preparation of this discussion was facilitated by a grant from the National Science Foundation

It seems to me, and, I believe, to Gagne, that if the relationship between learning and problem solving is to bear fruit, at this stage of research it is because the former is a field in which there is available some knowledge of process, some theory and some methodology which is viable and upon which the latter can draw. Such a position invokes two assumptions. The fundamental assumption is that we do have some scientific lore accumulated from some or all of the varieties of S—R learning research (to be taken in its broadest sense). Since this issue has been so ably dealt with by other participants in this symposium, it needs no further commentary here. The other assumption is that the knowledge derived from this research is relevant to problem solving. It is this assumption which is under discussion at present.

For the sake of perspective, let us recognize at the outset that there are many ways in which learning may be related to problem solving. At one extreme there is the effort begun by Watson (1920) and Thorndike (1922), implemented by Hull (1930) and Miller and Dollard (1941), and exemplified in contemporary form by Goss (1961) and Bourne and Restle (1959), to reduce the 'higher mental processes' to a complex form of S—R learning. To the extent that Skinner (1957) addresses himself to this problem, he is also a member of this camp.

In this kind of analysis it is assumed that the same variables, and the same laws relating those variables, are operating in both kinds of behavior, but in the 'higher mental processes' previously acquired associations which enter into sequential relations with each other and with the present stimulus situation, must be taken into account. Although there is no necessary opposition to the possibility that full understanding of problem solving behavior will require the discovery of new variables or new relationships between old variables, there is no search for them either. The emphasis is on the extension of that which has already been uncovered in simple learning situations to more complex forms of behavior.

At the other extreme is the position, taken by the gestaltists to the effect that S—R learning is a curious laboratory phenomenon that bears little relationship to the process of problem solution. This position is vividly delineated by Wertheimer (1945) in his book 'Productive Thinking' where he distinguishes between two types of processes. The one with which he is primarily concerned 'focuses on developing structural insight, structural mastery, and meaningful learning in the real sense of the word' (p 202). The other process is associationistic which he describes (p 201) in this way:

> The extremes at the other end are cases in which the result, the solution, is brought about by sheer chance discovery or merely by a succession of blind trials, by sheer external recall, sheer reliance on blind repetition, by blind

> drill or by prompting There are many situations the nature of which fundamentally allows of nothing but blind proceeding and blind finding as, for instance in widely used experiments with mazes discrimination tasks, and problem boxes Here all the factors that might furnish some clue to reasonably directed behavior are carefully excluded by the experimenter Under these conditions no genius however great, could at first do anything but engage in blind trials success could occur only by chance and then be repeated—unless meanwhile the arbitrary set is changed arbitrarily by the experimenter

Note the gallantry that prompted the omission of the obvious adjective for this kind of experimenter Wertheimer added spice to the schism by pointing out that, 'The difference between the extreme " of these two approaches " concern not merely intellectual procedures, they involve deep differences in human attitude" (1945, p 201)

Oddly enough this humanistic position is exemplified in its contemporary form by theorists who draw upon the program of the electronic computer as a source of inspiration Miller, Galanter, and Pribram (1960), for example, have written that they prefer "alternatives to nickel in the-slot, stimulus response conceptions of man" (p 2) Newell and Simon (1961), modestly offer their methodology as the long awaited, if not overdue, alternative to what they plainly consider the sadly limited "method-oriented' approach of S—R theorists From this point of view, learning, at least of the S—R variety, has little to do with problem solving Consequently, understanding of this complex process requires a different language, methodology, and theory from those which arose from learning research Like the earlier Gestaltists they emphasize the central as opposed to the peripheral process and the ahistorical as opposed to the historical variables Their method is frankly introspectionist

Although Gagne s predisposition seems to be toward the S—R kind of theorizing and experimentation, nevetheless, his approach falls somewhere between the two extremes For instance, he offers the proposition that problem solving, since it involves a relatively permament change in behavior as a function of experience, is itself, a form of learning But as he pursues this line of thought he points out that the usual behavior measures for learning experiments such as the number of trials or errors are, for the most part, inappropriate for problem solving He mentions the important place historically given to the sudden solution of a problem which is usually of an all-or none nature and which generally occurs after a single presentation of the stimulus situation Put another way, he seems to be saying that, if the two phenomena are defined in terms of the operations ordinarily used to measure them, there is more to be gained by recognizing the differences than by stressing the similarities

Such behavioral differentiation does not mean that there need be an

abrupt schism between the methodology of learning and of problem solving. Our work on reversal shift (e.g., H. H. Kendler & T. S. Kendler, 1962) and inferential behavior in children (e.g., T. S. Kendler & H. H. Kendler, 1962), if it qualifies as problem solving research, exemplifies the possibilities of methodological continuity, since the procedures are extensions of discrimination and instrumental learning techniques. However, this work also exemplifies Gagne's differentiation between behavior that meets the criterion for a higher level process like concept formation (e.g., reversal shifts) or problem solution (e.g., direct inferential solutions) from behavior in the same test situation that better meets the criterion for simple S—R learning (e.g., nonreversal shifts and noninferential solutions, respectively.) However, as Gagne points out, for most problem solving research, the modifications of methodology are much greater. One of the consequences is that such investigations have to deal with variables like instructions and individual differences which can be, or at least have been, ignored or minimized in learning research. The main point, however, is that while Gagne starts out to seek for similarities, he winds up emphasizing the differences between learning and problem solving methodology. Unlike the Gestaltist, he takes his point of departure from learning research and nowhere does he attempt to limit this methodology to verbal introspection.

Gagne goes on to other significant departures. His analysis of the independent and dependent variables uses S—R language, but the stimulus and the response in problem solving are to be modified in important ways. The stimulus becomes an "indefinite identity," capable of specification in conceptual categories, rather than in terms of physical stimulus characteristics. This statement contains the seed for many long ontological discussions about whether all definitions, including operational ones, are not actually specified in conceptual categories, albeit in categories specified by the experimenter. But Gagne seems to agree with the Gestaltists that the significant stimulus in the problem solving situation must be based on the perceptual response made by the *S*, not by the experimenter.

In a similar vein, he emphasizes that the significant response is not the observable "button pushing" or "door-opening" but rather some implicit response which may be generated by a class of situations and can engender a variety of overt behaviors. In other words, he emphasizes the central (or, at least, covert) process at the expense of the peripheral (or observable) response.

As to the historical ahistorical issue, Gagné leaves ample opportunity for the action of ahistorical variables, like the "patterning" of the situation as a function of instructions. But he does not neglect the contribution made by the reinforcement history of the problem solver which enters into his analysis in several ways. It appears when he discusses the role

of instructions in producing the recall of subordinate capabilities which are a composite of aptitude and training It appears when he discusses the thorny problem of individual differences It appears in its most systematic form, however, in the theory he offers about the ordering of human learning (p 312, Table 1) This is, basically, a construction that has the conditioned response at its foundation and problem solution at its pinnacle For several levels the conditioned response is not only the foundation, it also serves as the unit of construction, the building block, as it were Up through concept learning increasingly complex behavior simply consists of more or different S—R units (explicit and implicit) in various kinds of combination But as we approach the pinnacle, the unit of construction changes Herein lies the crux of the problem—are these latter stages simply still more complex combinations of S—R units or do new entities emerge here which require other kinds of analyses?

The Question of Emergent Properties

If S–R is the basic unit it would seem reasonable to pursue our understanding of the complex process by reducing it to its essential and then studying it in the simplest, most controlled way—by the study of the laws determining S–R association or, perhaps more pertinently, by studying how such S–R units interact in the simplest possible circumstances On the other hand if problem solution consists of something more or different than these elementary units, only a research strategy that deals with the complex process would be in order

Actually such a question can be raised at any level above the foundation That is, do new variables and new interrelationships emerge at each complexity level? Do new properties appear that cannot be predicted from the elements of which they are composed? Consider, for example, the first transition suggested by Gagne Here one can question whether all S—R units can be used equally readily to establish chains Perhaps there are some kinds of responses or perhaps some degrees of response strengths which form chains more easily than others If this should be the case, these properties would not be discovered until response chaining was investigated Studying simple conditioning by itself would not be sufficient As another example take the next level verbal learning Here one can inquire about whether the attachment of verbal labels functions in an analogous manner to any other response or does the verbal response have special properties which lend themselves to the development of longer chains than other responses Another relevant question is whether verbal links become covert more readily than other motor responses At the fourth level that of concept learning one can seek laws that specify how parallel series of

chains become organized into a concept Note that at the level of principle learning Gagne drops the S—R language It is not altogether clear whether this is done for the sake of convenient presentation or because, from here on, his theoretical constructs change status

The questions raised about the emergent characteristics of each level will ultimately require empirical answers But meanwhile those who are interested in problem solving must choose a level at which to study the process It seems reasonable that the more one is committed to the continuity between simple S—R learning and problem solving, and the less one is interested in coping with the enormous complexities of experimental control and manipulation in complex problem solution, the closer to the foundation one is likely to start Gagne, although he presents such an interesting formulation of this continuity and avers a commitment to it, nevertheless courageously chooses to emphasize the highest levels in his discussion (and research) This is, presumably, because he assumes that there are properties which emerge at this level that require study in their own right I would like to think that he also assumes that research can and should go on simultaneously at each of these levels, in order to produce data and subtheories that will provide the basis for one grand theoretical design capable of incorporating all levels, Such an ambitious goal may not be achievable and such a tolerant attitude may not be fruitful, but who can deny such a worthy aspiration?

References

Bourne L E, Jr & Restle F Mathematical theory of concept identification *Psychol Rev* 1959 66, 278–296

Goss A E Verbal mediating responses and concept formation *Psychol Rev* 1961, 68, 248–274

Hull C L Knowledge and purpose as habit mechanisms *Psychol Rev* 1930 37, 511–525

Kendler H H & Kendler T S Vertical and horizontal processes in problem solving *Psychol Rev*, 1962 69, 1–16

Kendler T S & Kendler, H H Inferential behavior in children as a function of age and subgoal constancy *J exp Psychol* 1962 64, 460–466

Miller G A, Galanter E & Pribram K H *Plans and the structure of behavior* New York Holt 1960

Miller N E & Dollard J *Social learning and imitation* New Haven Yale University Press 1941

Newell, A, & Simon H A Computer simulation of human thinking *Science* 1961 134, 2011–2017

Skinner B F *Verbal behavior* New York Appleton-Century-Crofts 19[illegible]

Thorndike, E L *Psychology of arithmetic* New York Macmillan 1922.

Watson, J B Is thinking merely the action of language mechanisms? *Brit J Psychol* 1920 11, 87–104

Wertheimer, M *Productive thinking* New York Harper, 1945

The Taxonomy of Human Learning: Overview

ARTHUR W MELTON
University of Michigan

These insightful papers on the validity and internal structure of the traditional categories of human learning processes represent the first concerted effort to examine these important systematic issues in the science of human learning It was not expected that the outcome would be a new, concise, neatly organized taxonomy of human learning processes and their subvarieties Rather, the expectation was that these efforts would give some impetus to more widespread, systematic consideration of taxonomic problems by those engaged in experimentation and theory within the area If such beneficial consequences of this symposium are realized, they will surely be traceable to the ideas presented in the preceding seven chapters Nevertheless, there may be some value in an attempt to examine in general terms the role and character of a taxonomy of human learning processes and to assess the progress in thinking about the problem that seems to have been made in this symposium

Why the Taxonomy of Human Learning?

All areas of psychological science have a taxonomic problem with respect to the phenomena investigated, but those that are defined by the operation of some function—such as learning, motivation, perception, performance—have no guidance from the structure of the organism in the generation of a "natural schema They have as their only resource the ingenuity of the scientist in discovering commonalities between apparently different phenomena and differences between apparently identical phenomena In view of the fact that the learning processes which are our principal concern are influenced by and undoubtedly interact with the motivational, perceptual, and performance processes—and our categories of learning should therefore interact with the categorizations of these other functions—it must be self-evident that any unilateral attempt to improve the taxonomy of learning processes is but a piece of the problem

of behavior science and one unlikely to be handled adequately without parallel efforts by psychologists concerned principally with the other functional variables

In view of the already limited taxonomic objective, the reason for further restricting our concern to human learning requires some discussion In particular, one might ask why there is the distinction between the taxonomic problem in all organismic learning and in human learning Aside from the argument that the problem needed to be brought into focus on a manageable segment of the whole problem, there are more substantive reasons for emphasizing *human* learning The first is that human learning encompasses the varieties of processes observable in animal learning, but goes beyond them, perhaps largely because of the capacities for verbal modes of symbolic, mediating' behavior in human beings This reason was validated by the discussions of the categories of human learning, in that in every category it was necessary to emphasize "mediating" processes and conceptual controls of the learning process These are interpreted, or interpretable, as verbal mediators and controls At the same time, it should be noted that our emphasis on human learning did not discourage Professors Grant and Kimble from broadening the taxonomic problem to include animal learning where the latter was most relevant to the discussion, i e , in the case of classical, operant, and instrumental conditioning

Another reason for emphasis on human learning deserves mention, especially since the forces of current history in scientific psychology sometimes fail to be explicitly recognized either because they are unknown or are taken for granted Within the past ten years there has been a marked resurgence of research and theory about human learning, as contrasted with the emphasis on animal learning, especially in the rat, in the decade of the 1940 s and early 1950 s The reasons for this resurgence are undoubtedly complex and manifold, and may well be a suitable topic for a doctoral thesis in the history of psychology early in the twenty-first century, but two factors in this trend seem quite clear During the late 1940's and early 1950 s the military, other governmental, and industrial agencies recognized the need for fundamental knowledge about human performance capabilities and limitations and a parallel need for fundamental knowledge about the optimal procedures for training men to meet performance requirements The tasks of concern to these extra psychological agencies were as varied as those of simple discriminative reaction to occasional signals as in vigilance studies through various forms of skill learning and knowledge acquisition to the problem solving involved in malfunction diagnosis in complex electronic equipment

These needs for a technology of human learning not only stimulated

fundamental research on human learning and performance, the effects of which continue to be in an accelerative phase, but also highlighted the issues surrounding the taxonomy of learning processes (e g , Cotterman, 1959) When one is confronted with a decision to use massed or distributed practice, to insist on information feedback or not to insist on it, to arrange training so as to maximize or minimize requirements for contiguous stimulus differentiation, etc , and discovers that the guidance received from experimental research and theory is different for rote learning, for skill learning, and for problem solving, taxonomic issues become critical and taxonomic ambiguities become frustrating, to say the least

The linkage of the problems of human learning and the problems of human performance in this resurgence of the psychology of human behavior is of considerable moment for the taxonomic problem One may properly speculate that the taxonomy of human performance implies a taxonomy of human learning processes, and vice versa The reason is that it is unlikely that one can make an appropriate prediction or assessment of human performance in any task situation without considering the historical antecedents of that performance capability, i e , without considering the characteristics of the learning processes that produce that performance capability This is true even though the performance has been brought to asymptote (if there is such) on the practice dimension Space does not permit the elaboration of this point, but some feel for its validity may be obtained from consideration of the possible differences between "asymptotic" human performance in a task that requires a simple quick reaction to a stimulus where the S–R relationship is one that involves a strong population stereotype (initial S–R compatibility) and a simple reaction time to a stimulus where the S–R relationship is one that involves the reverse of a strong population stereotype (initial S–R incompatibility) Even though both habits have been brought to identical asymptotic performance levels, according to some measure, it may be confidently predicted that the two learning histories involved in these habits—one involving strengthening and shaping of an existing dominant S–R relationship and the other involving the establishment of an originally nondominant response as a dominant response—will be revealed in some performance measures For example, performance measures under situational stress, or the retention of the habit over long intervals of no practice may reflect these different learning histories If the process route to a performance capability is latent in such capabilities even though some measures in some or many situations fail to reveal them, then human performance theory needs an adequate, referenceable taxonomy of human learning processes The need of human learning theory for an adequate taxonomy of human performance, and of human performance measures, is also readily understandable

in view of the fact that all measures of human learning are measures of performance change as a function of experience or practice

These, then, were some of the principal reasons for restricting our attention to the "narrower" taxonomic problems of human learning rather than the broader problems of learning in general or behavior in general Of course, the reader will also have observed that our concern is even narrower than these considerations forced it to be We have restricted our attention to the categories of human learning that have been brought under controlled laboratory observation, and we have not even attempted to consider all of these Perceptual learning, discrimination learning, and perhaps some forms of attitudinal and emotional learning have not been given explicit consideration as categories

Development of a Taxonomy

The natural history of taxonomies within a science, and within the science of human learning in particular, deserves some discussion as background for later comments

The noting of the similarities and differences of things and events is the first step in organizing knowledge about nature These observations are then the basis for classifications of things and events and the formulation of criteria of inclusion and exclusion This is essential to the generality that is the goal of a science, as well as to efficient communication among scientists It is for this reason that most of the categories of human learning that are included in the chapter titles of this volume have been with us since the beginning of experimental psychology ("probability" learning is the only truly recent newcomer)

These primitive categories are based on a sorting of learning processes into classes that have obvious differences at the descriptive level and they lean heavily on what may be called operational or quasi-operational criteria, and so may be called the classes of a primitive operational taxonomy Once formed, these primitive operational categories undergo a variety of changes as the scientific analysis and understanding of the phenomena progresses In short, a taxonomy reflects the stages of development of a science

Perhaps the first important development in taxonomy following the invention of laboratory tasks and procedures for the investigation of the first few of these primitive categories of human learning was recognition of a need to limit the generalization of empirical findings to a category, or even to a subclass of a category, until there was evidence to support a wider generalization This development did not come easily or quickly It was only 30 years ago that Carr (1933) felt compelled to caution in-

vestigators about the pitfalls in the "quest for constants." He not only warned against generalization of findings across categories, but also within categories, and even within specific experimental procedures when some critical aspect of the procedure was permitted to vary. He was saying, of course, that the phenomena or relationships between variables that are observed in some experimental context, as defined by the task, procedures, and measures, should not be assumed to be true for other tasks, procedures, or measures until shown to be so. This radical empiricistic pluralism gained many adherents, especially as a consequence of the importation of operationalism in the late 1930's. In my opinion, this position is as valid today as it was in the 1930's, and recognition of the truth of it is the first step in the development of a scientific taxonomy of human learning processes.

A science cannot, of course, tolerate such a complete limitation of the generality of an experimental phenomenon to the particular context in which the phenomenon is observed. Levels of generality, both intra-category and inter category, must be achieved either through systematic empirical investigations which bridge boundaries within or between postulated categories or by theories which employ hypothetical constructs or intervening variables to reveal the presence of similarities and differences that are more fundamental than those obtained at the observational level. As we all know, neither of these avenues to the achievement of a comprehensive taxonomy of human learning processes has been followed very far. Our theories are still for the most part intra-category theories, which is to say that they are theories about conditioning, about rote learning, about short-term memory, about concept learning, about probability learning, about skill learning, about problem solving. And as such, they are frequently very closely tied to a limited set of experimental operations within the category. This has certainly been the case with conditioning, probability learning and rote learning, and perhaps of the others as well. The failure of psychologists to follow the other avenue of parallel systematic empirical investigations within and between categories is understandable, in view of the magnitude of effort that would be required for such a venture and the hope that such effort would find greater rewards if guided by well-formed theory. There are, of course, some empirical evidences of generality of phenomena within and across the primitive categories of human learning, but it is resistant to summarization. It is, however, not surprising that the participants in the present discussion avoided the systematic empirical approach to the mapping of the interrelatedness of the categories of human learning.

Even though neither of these approaches has provided the basis for the revision of the whole taxonomy of human learning, both have provided

many limited revisions of the primitive categories First among the ways in which these primitive categories have been changed is through the refinement of their defining operations, and second—and usually going along hand and hand with the first—is through the identification and differentiation of subclasses of the category Major primitive categories have been absorbed into others or major categories have been split into two An excellent example of the consolidation of categories is currently in progress with the encompassing of the category of "memory span" ("immediate memory," or "short term memory) under the category of rote learning, in view of the commonality of the operation of direct presentation—rather than discovery—of the to be learned responses Again, "insight learning," as a category, now seems to be properly encompassed by the category of problem solving " On the other hand, the fission of categories is well represented by the now traditional separation of classical and instrumental conditioning to which Professors Grant and Kimble have added some important new thoughts in this volume

A very complex and radical revision of the primitive categories is suggested by Professor Fitts, in the case of skill learning In his view, the learning processes usually subsumed under the category show such strong similarities to the processes of language learning and utilization, concept learning and utilization, and problem solving that he questions whether skill learning should be retained as a major primitive category This development suggests that a major revision of the primitive categories may occur in the near future—a revision based on a deeper understanding of the similarities in the processes involved in these various kinds of learning and proved by the greater analytic and predictive power of the revised taxonomy based on processes rather than on primitive operation distinctions This is, of course, what is hoped for, and expected, as our data and theory permit movement from a phenotypical taxonomy to a genotypical taxonomy, from observables to constructs, from a variety of special theories tied to specific experimental operations to a general theory

In addition to these revisions in the major categories there have been equally important taxonomic changes within categories New subcategories have been invented or discovered and the operational and process distinctions between old subcategories have been refined Perhaps reflecting the increasing intensity of research effort with the techniques of classical conditioning instrumental conditioning, and rote learning, there have been major refinements and extensions within these three categories in the past few years Thus, Professor Grant discusses four subcategories of "classical conditioning and eight subcategories of instrumental conditioning where Hilgard and Marquis (1940) had one and four subcategories, re-

spectively In the case of subcategories of rote verbal learning, five are identified and differentiated by Professor Underwood, and Professor Postman discusses at least four subcategories of short term memory tasks (memory span, running memory span, continuous paired associate memory, and memory for single items) that may be considered as belonging to the rote learning category In much the same way, probability learning, concept learning, skill learning, and problem solving have had their subcategories increased and the distinctions between them refined We are forced to the conclusion that much research effort, analytical ingenuity, and theoretical sophistication have been devoted to making distinctions where none existed before or to adding varieties that were not known, at least in the laboratory, 10 or 20 years ago

It is clear that psychologists must expect, and are getting, a progressive movement of the taxonomy of human learning processes away from a strictly operational base and toward a theoretical base in which inferred processes become the categories One may ask, therefore, whether the operational taxonomy will be supplanted by or supplemented by the theory-based taxonomy It seems to me that the process will be that of supplementation, rather than substitution We will have both the operational taxonomy based on ostensive characteristics of the learning task or characteristics of stimuli and responses, and we will have a taxonomy based on the characteristics of inferred processes and their compounding

There are several reasons why this should be so In the first place, the descriptive anchor for the inferred process taxonomy will be the operational taxonomy of learning tasks and much of what we call theory will be the mapping of the genotypic relations of the various forms of learning that are operationally differentiated on the basis of their task characteristics In the second place, the operational taxonomy must continue to serve as an analytic, descriptive tool of the technology of human learning It will continue to be the basis on which varieties of learning outside the laboratory are reduced to categories or subcategories as the first step in the determination of the applicability of empirical law or theoretical understanding to the prediction of phenomena or efficient conditions of learning in the case in question

subdomains within which our empirical findings are generalized with great caution, but still serving to denote the categories of learning tasks engaged in by human organisms

In summary, the argument is that we necessarily started our science of human learning with a primitive taxonomy of learning tasks, that these have been subjected to combination and fission, and that subcategories have been refined and new subcategories added as a consequence of analytic research and theory Theory may eventually lead to a completely different set of categories based on process or construct distinctions, but a sophisticated operational task taxonomy will continue to be necessary for the analytic, empirical activities of our science

The Validity of the Traditional Categories of Learning

The basic purpose of this series of papers was to examine the traditional categories of human learning from the vantage point of expertness in each category As repeatedly emphasized, this involved both the identification of similarities and differences between categories—with stress on similarities, in view of the traditional implication of category differences—and the identification of similarities and differences within a category—with stress on differences, in view of the traditional implication of category homogeneity We have found many similarities between categories and many dissimilarities, or fundamental distinctions, within categories What is the significance of this for the taxonomy of human learning processes?

The form of the above statement with respect to the traditional categories of human learning—conditioning, rote learning, probability learning skill learning, concept learning, problem solving—may appear to be a verbal contrivance to create a straw man at this stage in the development of a science of learning In one sense it is The sophisticated investigator and theorist in human learning has for many years stressed the overlap of certain categories and the necessary distinctions between forms of learning commonly included within a category Nevertheless, there are reasons for believing that the further consideration of the validity of the traditional categories of human learning is not the further beating of a dead horse

First among these reasons is the observation that our elementary textbooks in psychology, which are presumed to treat the fundamentals of the science are quite confused and confusing with respect to these traditional categories of learning While most of them make the distinction between classical and operant conditioning and the distinction between rote learning skill learning and problem solving these categories are func-

tioning chiefly as class names defined by pointing to some familiar laboratory or real-life examples Even when the point is made that examples from different categories have similarities that might be the basis for intercategory similarities of processes or effects of major parameters—as in the case of the involvement of discovery of the correct response in the problem box, temporally organized skills, and problem solving or the serial S–R character of rote serial learning and some skills—there is no attempt to show that such is the case It may, of course, also be noted that our textbooks of learning per se struggle with this taxonomic problem and have come up with no clear and comprehensive statement of the problem or of even a partial solution to it

It is suspected that this state of affairs in textbooks merely reflects the indecision about taxonomic matters among students of learning As noted earlier, our investigators and theorists have learned well the lesson of operationism and the advisability of framing theories of limited scope Investigators limit the generalization of their empirical findings to the very narrow class represented by their specific experimental operations, and they are prodded in this direction by editors if they should falter And theorists—as suggested by Estes (p 90)—either work within the narrow confines of a category, or even a variety within a category, or formulate a limited scope theory that is not tied to a particular category of learning, but is tied to a restricted range of variables The net result has been a paucity of systematic thinking and writing about the full range of human learning and the possible basis on which this variety could be more meaningfully and validly categorized in terms of processes, phenomena, or the effects of variables

The most certain conclusion that one can reach about the traditional categories is that they are not the proper categories for use in understanding human learning even though they may serve a useful denotative function The best that can be said for them is that each category does include a task aspect or behavior requirement which is important and is given heavy weight or emphasis in at least some of the subcategories within it Thus, classical conditioning and the various forms of rote verbal learning (with the exception of "verbal discrimination learning" but including the short-term memory methods) place no formal requirement on the learner to "discover" the correct response, instead, they supply the nominal stimulus and the to-be-learned response (or a reasonable facsimile thereof), and these are usually in the appropriate temporal relationship although the temporal arrangement is a critical parameter By contrast, operant conditioning, instrumental conditioning, and in fact all the remaining varieties of learning involve a process of "discovery" of the correct re-

sponse, i.e., the individual must emit the response upon which reinforcement is contingent. For simplicity, these two great classes of learning processes may be called *simple associative* and *selective*.

However, this distinction may be one of degree. Kimble (p. 41) argues cogently for the inclusion of classical conditioning in instrumental learning, and Jones (1962) has argued for the inclusion of instrumental learning in classical conditioning, at least in the later phases of it. Similarly, Underwood (pp. 57–73) has made a strong case for the presence of stimulus selection, response learning, concept learning, and mediation in rote verbal learning and all of these reflect learner-generated selective learning processes over and above the simple associative learning that is prescribed by the methodology. Again, simple associative learning, such as memory (short- or long-term) for events (stimuli, responses, reinforcement or "feedback") surely is involved in important ways in probability learning, concept learning, the varieties of skill learning, and problem solving.

Insistence that all of the categories of learning involve both simple associative learning and selective learning does not involve a commitment to a uniprocess or to a dual process theory of all learning. Instead, it merely specifies that, at this stage in our investigation of learning we have no *pure* technique for isolating these processes in the laboratory. At the same time, it may be recognized that the categories of selective learning vary greatly in amount and conditions of "discovery" of the to-be-learned response and that within each category the involvement of "discovery" may be independently manipulated. The dependence of the learner on simple associative memory processes may also be manipulated (Cahill & Hovland, 1960, Gagne, p. 308). It also seems likely that, following Underwood's illuminating discussion of the selective learning aspects of rote verbal learning, it may be possible to reduce the involvement of such unwanted aspects by further refinement of the rote learning methods or materials. In short, it appears that the refinement of observations on simple associative learning may provide closer approximations to the learning situation required for the determination of the laws of such learning, where "discovery" of the relevant stimulus and appropriate response is brought to near zero. Likewise, systematic investigation of the effects of variation in the number and strength of responses other than the one to be selectively reinforced when it is emitted may provide information about the "discovery" process in each of the relevant categories that makes them better described in terms of such a process than by the usual category names.

Other important process characteristics of human learning show only a limited correlation with the traditional categories. Prominent among

these is the discrete S–R relationship in contrast to the simultaneous or successive temporal organization of responses The former is represented in some classically conditioned responses, and it is sometimes thought to be represented in paired associate verbal learning Again the *pure* case is probably not attainable, but it may be suggested that one reason for the failure of "transition" experiments to reveal a close parallelism of the phenomena of classical conditioning and paired associate learning (Underwood, p 49) is that the latter characteristically involves the simultaneous establishment of a number of S–R relationships That is, paired associate learning has a 'list factor' in it that is not to be found in classical conditioning, and it is well known that this 'list factor' is important In list learning the interval between presentation of an S–R pair and the testing of the association is filled with other similar activity, this is important (Peterson, Saltzman, Hillner, & Land, 1962) perhaps because unlike the usual CR paradigm, it forces a high rate of 'deconditioning' between presentation and test of a pair Simple classical conditioning is, therefore, not necessarily directly comparable with verbal paired associate learning on this dimension of simplicity complexity of the association matrix But neither is it necessarily different from some forms of discrete S–R skill learning, as suggested by Grant (p 5–7) in his discussion of "anticipatory instructed conditioning" and as implied by Fitts (pp 253–261) in his analysis of the discrete S–R skill Also, there are strong similarities between paired associate verbal learning and continuous discrete S–R learning and performance normally considered as an example of skill learning perhaps again because of the ubiquitous involvement of verbal processes in all forms of adult human learning

Finally, the serial or temporal organization of responses is not the defining property of a traditional category of learning even though it is obviously a critical dimension of certain subcategories of skill learning and may be represented in serial rote learning It is found also in many varieties of instrumental learning and problem solving In all cases we are some distance removed from an understanding of such serial temporal organization of behavior, but it is clear that the analysis of the problem is leading to similar conceptualizations of the problem and of process characteristics of the behavior, regardless of traditional category differences In all instances there is agreement that the nature of the stimulus involved in serial acts is unclear and that the description of the processes involved almost certainly require intervening variables such as "mediating responses" which have a stimulus function and/or control of the sequencing of responses by "second order' habits (concepts rules schemata) (Fitts pp 270–276, Underwood, pp 60–62)

These are but several examples of the way in which intensive analyses

of the varieties of human learning have revealed more meaningful, and presumably more valid, process criteria for a taxonomy than those employed in our primitive operational taxonomy One distinguishing characteristic of these more valid criteria is that they refer to the process characteristics of the learning, rather than to a mish mash of procedural and topographic (i e , perceptual, motor, verbal, "central") criteria This process taxonomy does not, of course, imply the independence of processes and procedural or task characteristics, quite the contrary, it insists that there is variation in the involvement of a process as a function of the characteristics of the task, the nature of the stimuli and responses required, and the procedural controls of the learner's behavior The processes must therefore be thought of as variables or "dimensions" (Melton, 1941) of human learning that may combine in different ways in different instances of learning Perhaps some of these processes that are important in the description and understanding of human learning may be absent from particular instances of learning but when present they are subject to variation in amount or weight

One of the interesting suggestions derived from this direction of change in the conceptual taxonomy of human learning is that there may be very much more commonality of processes across the traditional categories of human learning than those categories would suggest This is, of course, what general theorists have assumed, but they have been unable to incorporate in their theories a sufficient variety of processes or phenomena to serve as guides for research on all the primitive operationally-defined categories This deficiency is, in turn, traceable in at least some instances to the assumption that the "simpler" forms of learning were free, or could be made free, of some of the processes that complicate other forms of learning It is, therefore, of some importance to note the extraordinary agreement among the participants in this symposium in identifying the critical issues in the further analysis of the categories of learning for which they were responsible

A first point of agreement was that there must be, in all types of learning better specification control and measurement of the prior history of the *S*, i e , of transfer of learning Pretesting the CS in conditioning experiments specification of letter sequence and word sequence habits in verbal learning, assessing encoding and recoding responses to stimuli in studies of short term memory and incidental learning identifying and measuring the strength of population stereotypes" in all types of skill learning and defining the source and availability of hypotheses in problem solving—all relate to this ubiquitous transfer variable In the case of probability learning and concept learning the problem of transfer became critical to conflicting theoretical approaches

A second point of agreement was that there may need to be a concept of "second-order" or "higher-order" habits in the analysis of all forms of learning, and at least there is this need in a number of otherwise diverse kinds of learning Second order habits (some would call them concepts) are said to be involved in verbal operant conditioning, in the "selector mechanism" and "stimulus selection" process in rote verbal learning, in the cognitive sets involved in various forms of learning, and of course in concept learning per se and in problem solving where the criterion of problem solving *is* the concept

A third point of agreement was the need for the more effective analysis and control of stimulus variables This is, of course, one of the criteria for differentiation of classical and operant or instrumental conditioning, it also became a major consideration in the distinction between the nominal and effective stimulus in rote verbal learning, it is a major consideration in probability learning where, according to Estes (p 95), it is important to set the *S* for reaction to discrete events rather than to series of events, and it is critical to the definition of "incidental learning" and to the results of experiments on it

A fourth point of somewhat general agreement, although not so explicitly considered, was the need for the concept of "mediating' processes Such mediators may, of course, be identified in some cases as 'second-order habits" or concepts, but there was frequent reference to—and certainly no aversion to—the employment of mediators as first-order habits that mediated the stimulus and response relationship of other first-order habits

If, as seems to be the case, further research on the processes and phenomena of the traditional categories of learning leads us to a taxonomy in terms of process dimensions, these four points of agreement should become involved in the resultant dimensional taxonomy

CONCLUSION

Perhaps the first and most valid conclusion from this review of the taxonomy of human learning is that any conclusion is sure to be inconclusive Twenty-five years ago I shared with a number of other psychologists a concern about the slow progress in the development of a taxonomy that would be an effective guide to the improvement of our understanding of learning I concluded that too much needed to be known before more than a rough sketch of a new taxonomy might be proposed (Melton, 1941) In spite of vigorous theoretical and experimental work in this interim period, the same conclusion seems to hold today But there are some differences

From all that has been said in this volume it should be clear that the traditional, primitive, categories of human learning which serve as chapter titles have little usefulness in the scientific analysis of human learning, even though they may still serve a useful purpose in the technology of human learning Empirical work and the limited theories during this period have made it amply clear that the most useful set of prime categories in any contemporary taxonomy is the rather large set, and steadily increasing set, of subcategories of those primitive major categories This set of new, operationally narrower categories would, for example, include the subcategories of classical and operant conditioning, the subcategories of rote verbal learning, etc

This new operational taxonomy would have several advantages It would reflect the way psychologists are going about the development of a better understanding of the whole range of human learning, both in the laboratory and in theory It would avoid the implication of similarity in different categories of learning, when in fact the dissimilarities—process-wise and phenomenon wise—are much more striking than the similarities In short, it would reduce the probability that psychologists or the users of psychological knowledge about human learning would be tricked by names But perhaps the most important consequences of this radical pluralistic taxonomy would be to encourage the movement of thought and research in the direction of a reintegration of the categories in a new, less pluralistic taxonomy

The ways in which this latter effect might be realized are many In the first place, the elimination of the old category names should facilitate the recognition of process similarities among kinds of learning heretofore thought of as members of different categories of learning, for example, subcategories of what are now thought of as instrumental conditioning, concept learning, and skill learning would be more readily recognized as having important, even dominating, process characteristics in common It would, in short, counteract the extreme specialization of research within the field of human learning, which has resulted in "specialists" in skill learning in rote learning, in classical conditioning, in instrumental conditioning etc

If this favorable effect were realized, and process similarities became the source of experimental and theoretical analyses that bridged the old category distinctions, another favorable outcome would surely be a marked increase in what Underwood (p 49) calls "transition experiments" And it may be predicted with confidence that attempts to perform transition experiments will increase the number of different varieties of learning with which we work in the laboratory, and such attempts will encourage the development of a dimensional approach, i e , a continuum approach, to

the process variables that are involved in learning tasks that are today considered different ' because of different category memberships

The technique of adopting again a radical pluralistic taxonomy, as a means of encouraging the recognition of new and more fundamental ways of ordering our knowledge about human learning may not be the only effective way to achieve this effect But it is essential that some way be found to do so An improved taxonomy and an improved understanding of the entire range of human learning requires a much greater emphasis on the determination of the effects of similar processes—and the variation in the characteristics of such similar processes—in otherwise different contexts of other variables This does not necessarily mean that less emphasis should be placed on the exploration of the process characteristics of learning in relatively homogeneous contexts of other variables It is to be hoped that both approaches to an understanding of human learning will find adequate and increasing support and encouragement

References

CAHILL H E & HOVLAND C I The role of memory in the acquisition of concepts *J exp Psychol* 1960 59, 137–144

CARR, H A The quest for constants *Psychol Rev* 1933 40 514–522

COTTERMAN T E Task classification An approach to partially ordering information on human learning *WADC Tech Note* 58–374 January 1959

HILGARD E R & MARQUIS D G *Conditioning and learning* New York Appleton Century 1940

JONES J E Contiguity and reinforcement in relation to CS UCS intervals in classical aversive conditioning *Psychol Rev* 1962 69 176–186

MELTON A W Learning In W S MONROE (Ed) *Encyclopedia of educational research* New York Macmillan 1941 Pp 667 686

PETERSON L R SALTZMAN D HILLNER K & LAND V Recency and frequency in paired associate learning *J exp Psychol* 1962 63 396–403

Author Index

The numbers in italics indicate the pages on which the complete references are listed

N

O

P

Subject Index

S